

( Translation )

PATENT OFFICE  
JAPANESE GOVERNMENT



This is to certify that the annexed is a true copy of  
the following application as filed with this Office.

Date of Application: March 14, 2001

Application Number: Japanese Patent Application  
No. 071769/2001

Applicant(s):  
- Kousaku OKUBO  
- Hitachi Software Engineering Co., Ltd.

October 26, 2001

Commissioner,  
Patent Office

Kozo OIKAWA (seal)

Certificate No. 2001-3093456

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

JC971 U.S. PTO  
10/077662  
02/14/02

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日  
Date of Application:

2001年 3月14日

出 願 番 号  
Application Number:

特願2001-071769

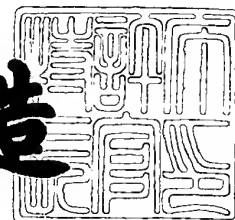
出 願 人  
Applicant(s):

大久保 公策  
日立ソフトウェアエンジニアリング株式会社

2001年10月26日

特 許 庁 長 官  
Commissioner,  
Japan Patent Office

及 川 耕 造



出証番号 出証特2001-3093456

【書類名】 特許願  
【整理番号】 12B025  
【提出日】 平成13年 3月14日  
【あて先】 特許庁長官 殿  
【国際特許分類】 G06F 17/30  
【発明の名称】 知識データベース及び知識データベースの構築方法  
【請求項の数】 29

【発明者】

【住所又は居所】 大阪府箕面市瀬川2丁目11番26号

【氏名】 大久保 公策

【発明者】

【住所又は居所】 神奈川県横浜市中区尾上町6丁目81番地 日立ソフト  
ウェアエンジニアリング株式会社内

【氏名】 田村 卓郎

【発明者】

【住所又は居所】 神奈川県横浜市中区尾上町6丁目81番地 日立ソフト  
ウェアエンジニアリング株式会社内

【氏名】 山下 巖

【特許出願人】

【住所又は居所】 大阪府箕面市瀬川2丁目11番26号

【氏名又は名称】 大久保 公策

【特許出願人】

【識別番号】 000233055

【氏名又は名称】 日立ソフトウェアエンジニアリング株式会社

【代理人】

【識別番号】 100091096

【弁理士】

【氏名又は名称】 平木 祐輔

特 2 0 0 1 - 0 7 1 7 6 9

【選任した代理人】

【識別番号】 100102576

【弁理士】

【氏名又は名称】 渡辺 敏章

【手数料の表示】

【予納台帳番号】 015244

【納付金額】 21,000円

▼ 【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9722155

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 知識データベース及び知識データベースの構築方法

【特許請求の範囲】

【請求項 1】 自然言語情報に含まれる句をオブジェクト値として有する複数の句オブジェクトと、前記句オブジェクト間の関係を定量的に表現した値をリレーション値として有する前記句オブジェクト間のリレーションとを蓄積したことを特徴とする知識データベース。

【請求項 2】 DNA配列又は蛋白質配列をオブジェクト値として有する複数のオブジェクトと、前記DNA配列間又は前記蛋白質配列間の関係を相同性計算あるいは実験的な測定により求めて定量的に表現した値をリレーション値として有する前記オブジェクト間のリレーションとを蓄積したことを特徴とする知識データベース。

【請求項 3】 複数のオブジェクトと、前記オブジェクト間のリレーションとを蓄積した知識データベースであって、

前記オブジェクトは、DNA配列をオブジェクト値として有するDNA配列オブジェクト又は蛋白質配列をオブジェクト値として有する蛋白質配列オブジェクトと、前記DNA配列又は蛋白質配列に関連付けられた自然言語情報に含まれる句をオブジェクト値として有する句オブジェクトとを含み、

前記DNA配列オブジェクト又は蛋白質配列オブジェクトと前記句オブジェクトとの間のリレーションは予め定義されたりレーション値を有することを特徴とする知識データベース。

【請求項 4】 画像をオブジェクト値として有する複数のオブジェクトと、前記オブジェクト間の関係を定量的に表現した値をリレーション値として有する前記複数のオブジェクト間のリレーションとを蓄積したことを特徴とする知識データベース。

【請求項 5】 複数のオブジェクトと、前記オブジェクト間のリレーションとを蓄積した知識データベースであって、

前記オブジェクトは画像をオブジェクト値として有する画像オブジェクトと、前記画像に関連付けられた自然言語情報に含まれる句をオブジェクト値として有

する句オブジェクトとを含み、

前記画像オブジェクトと前記句オブジェクトとの間のリレーションはそれぞれ予め定義されたりレーション値を有することを特徴とする知識データベース。

【請求項 6】 複数のオブジェクトと、前記オブジェクト間のリレーションとを蓄積した知識データベースであって、

前記オブジェクトは医学分野及び／又は生物学分野で用いられる用語を含む句をオブジェクト値として有する句オブジェクトを含み、前記リレーションは前記句オブジェクトの間の関係を定量的に表現した値をリレーション値として有することを特徴とする知識データベース。

【請求項 7】 請求項 6 記載の知識データベースにおいて、同一のオブジェクトと見なすべき複数のオブジェクトをシノニムオブジェクトグループとして登録したシノニムオブジェクト辞書を備えることを特徴とする知識データベース。

【請求項 8】 請求項 6 記載の知識データベースにおいて、前記オブジェクトは DNA 配列をオブジェクト値として有する DNA 配列オブジェクト、蛋白質配列をオブジェクト値として有する蛋白質配列オブジェクト、又は画像をオブジェクト値として有する画像オブジェクトを含み、前記 DNA 配列オブジェクト、蛋白質配列オブジェクト又は画像オブジェクトと前記句オブジェクトとの間のリレーションは予め定義されたりレーション値を有することを特徴とする知識データベース。

【請求項 9】 請求項 8 記載の知識データベースにおいて、相同性が高い DNA 配列をオブジェクトとして有する複数の DNA 配列オブジェクト、あるいは相同性が高い蛋白質配列をオブジェクト値として有する複数の蛋白質配列オブジェクトを同一オブジェクトグループとして登録したシノニムオブジェクト辞書を備えることを特徴とする知識データベース。

【請求項 10】 請求項 1 又は 6 記載の知識データベースにおいて、前記句オブジェクトは自然言語で表現された知識（言語化知識）中の句をオブジェクト値として有し、それぞれの句が前記言語化知識内で近傍に存在する頻度を定量化した値を対応する句オブジェクト間のリレーションのリレーション値として有することを特徴とする知識データベース。

【請求項 1 1】 請求項 1 又は 6 記載の知識データベースにおいて、前記句オブジェクトは書籍の索引に含まれる句をオブジェクト値として有し、2つの句が当該書籍の同じ頁あるいは同じ段落に存在する頻度を定量化した値を対応する2つの句オブジェクト間のリレーションのリレーション値として有することを特徴とする知識データベース。

【請求項 1 2】 請求項 1 又は 6 記載の知識データベースにおいて、前記句オブジェクト間のリレーションは、それぞれの句が、自然言語で表現された知識の知識単位において存在するパターンの類似性（存在プロファイル）を定量化した値をリレーション値として有することを特徴とする知識データベース。

【請求項 1 3】 医学分野及び／又は生物学分野で用いられる用語を含む句をオブジェクト値として有するオブジェクトと、前記オブジェクト相互間の関係を定量的に表現した値をリレーション値として有するオブジェクト間のリレーションとを含む知識データベースから、1又は複数のキーワードに対応するオブジェクト（クエリーオブジェクト）に関連するオブジェクトを前記オブジェクト間のリレーションを元に抽出し、前記クエリーオブジェクト及び抽出されたオブジェクトをその間のリレーションと共に表示することを特徴とする方法。

【請求項 1 4】 請求項 1 3 記載の方法において、前記クエリーオブジェクトをオブジェクト値の入力又は検索により選択し、前記選択されたクエリーオブジェクト及び抽出されたオブジェクト並びにその間のリレーションをリスト表示又はグラフィカル表示することを特徴とする方法。

【請求項 1 5】 請求項 1 3 記載の方法において、前記クエリーオブジェクトに対して、強いリレーションを持つオブジェクトを階層的に求め、前記クエリーオブジェクト及び階層的に求めたオブジェクト並びにその間のリレーションをリスト表示又はグラフィカル表示することを特徴とする方法。

【請求項 1 6】 自然言語により表現された医学分野及び／又は生物学分野の知識（言語化知識）より、その表現において使用されている句を抽出し、その句をオブジェクト値として有するオブジェクトを生成するとともに、各オブジェクト間の関係を定量的に表現した値をリレーション値として有するオブジェクト間のリレーションを生成し、生成したオブジェクトとオブジェクト間のリレーシ

ョンとを蓄積することを特徴とする知識データベースの構築方法。

【請求項 1 7】 請求項 1 6 記載の知識データベースの構築方法において、前記言語化知識を、予め用意した 1 文字又は複数文字より成る文分離文字列により文へ分解し、続いて、予め用意した 1 文字又は複数文字より成る句分離文字列により句に分解し、これを句オブジェクトのオブジェクト値とし、2つの句が前記言語化知識内で近傍に存在する頻度を対応する2つの句オブジェクト間のリレーションのリレーション値とすることを特徴とする知識データベースの構築方法。

【請求項 1 8】 請求項 1 6 又は 1 7 記載の知識データベースの構築方法において、前記言語化知識として書籍の索引、書籍の目次、学术论文のタイトル、書籍の本文、学术论文の本文、及び／又はWebページの本文を使用することを特徴とする知識データベースの構築方法。

【請求項 1 9】 請求項 1 6 記載の知識データベースの構築方法において、句オブジェクトのマスターデータとして書籍の索引に含まれる句を用い、2つの句が前記書籍の同じ頁に存在する頻度を対応する2つのオブジェクト間のリレーションのリレーション値とすることを特徴とする知識データベースの構築方法。

【請求項 2 0】 請求項 1 6 記載の知識データベースの構築方法において、句オブジェクトのマスターデータとして予め指定した任意の句を用い、2つの句が前記言語化知識内で近傍に存在する頻度を対応する2つのオブジェクト間のリレーションのリレーション値とすることを特徴とする知識データベースの構築方法。

【請求項 2 1】 請求項 1 6 記載の知識データベースの構築方法において、前記言語化知識として学术论文を用い、学术论文のタイトル及び／又は本文を含む情報をネットワーク経由等により定期的を取得し、取得した情報から句オブジェクト及びリレーションを抽出し、知識データベースをアップデートすることを特徴とする知識データベースの構築方法。

【請求項 2 2】 請求項 1 6 記載の知識データベースの構築方法において、同一のオブジェクトと見なす複数のオブジェクト及びそれらに関連したリレーションを併合することを特徴とする知識データベースの構築方法。

【請求項 2 3】 DNA配列を含む情報からDNA配列を抽出し当該DNA配列をオブジェクト値とするDNA配列オブジェクトを生成し、2つのDNA配列オブジェクト間に、対応する2つのDNA配列間の関係を相同性計算あるいは実験的な測定により求められた定量的な値をリレーション値として有するリレーションを生成し、生成したDNA配列オブジェクト及びDNA配列オブジェクト間のリレーションを蓄積することを特徴とする知識データベースの構築方法。

【請求項 2 4】 請求項 2 3 記載の知識データベースの構築方法において、前記DNA配列に関連付けられた自然言語情報に含まれる句を抽出しそれをオブジェクト値とするオブジェクトを生成し、句オブジェクトと対応するDNA配列オブジェクトの間に定義されたりレーション値を有するリレーションを設定することを特徴とする知識データベースの構築方法。

【請求項 2 5】 請求項 2 4 記載の知識データベースの構築方法において、DNA配列とそれに関係付けられた自然言語情報を含む情報をネットワーク経由等により定期的を取得し、前記取得した情報からDNA配列オブジェクト、句オブジェクト及びリレーションを抽出し、知識データベースをアップデートすることを特徴とする知識データベースの構築方法。

【請求項 2 6】 請求項 2 3 記載の知識データベースの構築方法において、相同性が高いDNA配列をオブジェクト値として有する複数のDNA配列オブジェクトを同一オブジェクトと見なし、同一オブジェクトと見なした複数のDNA配列オブジェクト及びそれらに関連したりレーションを併合することを特徴とする知識データベースの構築方法。

【請求項 2 7】 自然言語により表現された句をオブジェクト値として有する複数のオブジェクトと、前記オブジェクト間の関係を定量的に表現した値をリレーション値として有する前記複数のオブジェクト間のリレーションとを蓄積した第1及び第2の知識データベース間で演算を行い、第3の知識データベースを生成することを特徴とする知識データベースの生成方法。

【請求項 2 8】 自然言語により表現された句をオブジェクト値として有する複数のオブジェクトと、前記オブジェクト間の関係を定量的に表現した値をリレーション値として有する前記複数のオブジェクト間のリレーションとを蓄積し

た第 1 及び第 2 の知識データベースを用い、前記第 1 及び第 2 の知識データベースにおける注目するリレーションのリレーション値の違いを比較するグラフを表示することを特徴とする方法。

【請求項 29】 請求項 27 又は 28 記載の方法において、複数の知識データベース間で、各知識データベースに含まれる句オブジェクト間のリレーションのリレーション値を標準化することを特徴とする方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、医学の分野や生物学の分野においてさまざまな形で蓄積された知識を、その知識を構成する‘物質名や事象’（オブジェクト）とそれぞれ相互の関係（リレーション）として再構築し、知識の構造化を可能とする技術に関する。

【0002】

【従来の技術】

近年の医学・生物学研究の隆盛に伴って大量の研究成果が報告され、学術論文、教科書、記事などの文章として、また、DNAの塩基配列（DNA配列）、蛋白質のアミノ酸配列（蛋白質配列）、三次元コーディネートといったデータとして、蓄積されてきている。この情報量は、急激に増加を続けており、教科書はますます厚く、また、報告されるDNA・蛋白質配列の量は、蓄積量が2倍になるのに1年かからない状況である。

【0003】

こうした大量の医学・生物学情報を蓄積し、情報へのアクセスを実現している例としては、米国国立医学図書館（NLM）の国立バイオテクノロジー情報センター（NCBI）が運用するPubMedデータベースがあげられる（<http://www.ncbi.nlm.nih.gov/>）。PubMedにおいては、医学・生物学に関連した膨大な量の学術論文の情報をデータベース化しており、それらの論文情報へ、タイトルや要旨に含まれる語句、著者名、雑誌名、発行年度、関連DNA・蛋白質配列情報、関連論文などからアクセスすることができる。

【0004】

また、その他の例として、Weizmann Institute of ScienceのMichael Rebhan等によるGeneCardsがある(<http://bioinformatics.weizmann.ac.il/cards/>)。インターネット上のパブリックデータベースより、自動的に遺伝子に関連した情報を集め、個々の遺伝子を単位として情報を纏める仕組みを用意しており、目的の遺伝子に関する最新且つ網羅的な情報の提供を実現している。

【0005】

【発明が解決しようとする課題】

多くの学術分野においては、研究の進展は真理の発見に結びつき、その時点で、誤りが判明した仮説に結びついていた大量の知識が破棄されることで情報の整理が行われ、その上に立って、次の研究が展開されて行くものと考えられる。然るに、医学・生物学の分野においては、こうした情報の整理がなされないまま、断片的な知識の蓄積が進み、情報の増殖が進んでいるものと考えられる。また、多くの研究者が自らの専門分野について知識を掘り下げる結果、研究が専門化・細分化していることも、横断的な情報から真理を求めることを困難としている。

【0006】

従来の技術においては、学術論文や書籍、Webページといった知識単位毎に情報が管理されており、知識単位を越えた情報を得るには、たとえば、複数の知識単位から論理的に導かれる知識を得るためには、それに必要な知識単位を入手し、それらの内容を理解し、記憶し、頭の中で結び付ける必要があった。

【0007】

これは、従来の技術においては、知識単位を蓄積することはできるものの、知識単位を構成する知識要素の間の構造や関連（階層構造、関係構造）を情報として持ち、利用することができなかったことを意味している。同じ理由により、自然言語により表現された医学・生物学分野の知識（言語化知識）、DNA・蛋白質配列情報など、異なる種類の情報より、論理的に知識を導き出すことはできなかった。

【0008】

知識要素の関連を扱う仕組みを実現し、医学・生物学の分野に応用することで

、解剖学、臨床、生化学といった専門分野を越えた知識を結合し、多面的な知識の観察や潜在的な知識の発掘をすることが可能となり、また、医学・生物学研究の過程で生じる大量の情報より、必要な情報を引き出す手段を提供することが可能となる。また、知識の図形表現による内容理解の補助、ある知識を説明する別表現の発見、知識の階層構造の発見や、異なる表現による同一知識の検出などが可能となる。本発明は、このようなことを可能にする手段を提供することを目的とする。

【 0 0 0 9 】

【課題を解決するための手段】

本発明では、前記目的を達成すべく、対象となる知識を、知識要素であり種類に応じた1つの値を持つ「オブジェクト」と、オブジェクト間の関連として方向性のない1つの値を持つ「リレーション」によって定義することにより、知識の合併、抽出、図示などの情報処理を可能とし、また、それを応用した知識情報処理を実現する。

【 0 0 1 0 】

具体的には、オブジェクトの種類として、自然言語により表現された知識、特に医学・生物学分野の知識（言語化知識）を対象として、その知識を構成する専門用語や熟語、文節などの句を知識要素とした「句オブジェクト」や、DNA配列や蛋白質配列を対象としてDNA配列や蛋白質配列を知識要素とした「DNA配列オブジェクト」「蛋白質配列オブジェクト」、画像を知識要素とした「画像オブジェクト」などを定義する。たとえば、句オブジェクトにおいては、知識のリソースである言語化知識において2つの句が近傍に存在する頻度によりオブジェクト間のリレーション値を設定する。また、DNA配列オブジェクト及び蛋白質配列オブジェクトにおいては、DNA配列及び蛋白質配列間の相同性スコア（Smith, T. F. & Waterman, M. F. 1990, Proc. Nat. Acad. Sci., U.S.A. 87, 118-122）や最大一致長とその一致率によって求めた値をもとにオブジェクト間のリレーション値を設定する。また、画像オブジェクトにおいては、画像のパターンマッチングのスコア、画像の特性値（明度、粒度等）の類似スコアや、画像の取得状況を元に設定した値をもとにオブジェクト間のリレーション値を設定する。

【0 0 1 1】

ここで、複数のオブジェクトと全オブジェクト間のリレーションからなる1つの知識情報を「知識データベース」とするとき、オブジェクトは知識データベース内でユニークな値を持つものとする。それにより、2つの知識データベースを合併する際、同一の値を持つオブジェクトを介して、2つの知識集合を融合することが可能となる。たとえば、2つの異なる専門分野の情報源（知識リソース）から作成した知識データベースを合併することにより、2つの専門分野の知識が融合した知識データベースを生成することが可能となる。

【0 0 1 2】

リレーション値の算出方式の定義は、知識データベース利用の目的に応じて設定可能とする。たとえば、DNA配列オブジェクトや蛋白質配列オブジェクト間のリレーションを、DNA配列や蛋白質配列の相同性によって設定し、物理的な類似性の情報として利用したり、また、遺伝子発現プロファイル測定実験の結果に基づいて設定し、分子生物学的な性質の情報として利用したりすることを可能とする。

【0 0 1 3】

リレーションの値は、「関係の大小」及び「対象となる2つのオブジェクトを同一とみなす」ことが表現可能であることとする。これにより、異なる種類のオブジェクト間で、ある一組のオブジェクトを同一とみなす表現が可能となる。

【0 0 1 4】

ここで、前記目的を達成するための本発明による手段を以下に示す。

(1) 自然言語情報に含まれる句をオブジェクト値として有する複数の句オブジェクトと、句オブジェクト間の関係を定量的に表現した値をリレーション値として有する句オブジェクト間のリレーションとを蓄積したことを特徴とする知識データベース。

句には、用語、熟語、文節などが含まれる。

【0 0 1 5】

(2) DNA配列又は蛋白質配列をオブジェクト値として有する複数のオブジェクトと、DNA配列間又は蛋白質配列間の関係を相同性計算あるいは実験的な測定に

より求めて定量的に表現した値をリレーション値として有するオブジェクト間のリレーションとを蓄積したことを特徴とする知識データベース。

【 0 0 1 6 】

(3) 複数のオブジェクトと、オブジェクト間のリレーションとを蓄積した知識データベースであって、オブジェクトは、DNA配列をオブジェクト値として有するDNA配列オブジェクト又は蛋白質配列をオブジェクト値として有する蛋白質配列オブジェクトと、DNA配列又は蛋白質配列に関連付けられた自然言語情報に含まれる句をオブジェクト値として有する句オブジェクトとを含み、DNA配列オブジェクト又は蛋白質配列オブジェクトと句オブジェクトとの間のリレーションは予め定義されたりレーション値を有することを特徴とする知識データベース。

DNA配列又は蛋白質配列に関連付けられた自然言語情報に含まれる句には、DNA配列又は蛋白質配列を特定するための識別子（アクセッション番号、名称、記号、符号、識別番号等）も含まれる。

【 0 0 1 7 】

ある2つのオブジェクトが同じ種類であっても異なる種類であっても、それら2つのオブジェクトの種類に応じたりレーション値の算出方式を定義可能とすることにより、種類の異なるオブジェクトの組み合わせによる知識の表現が可能になる。また、同じ種類のオブジェクトに対しても、目的に応じたりレーション値の算出方式を定義可能とすることにより、目的に適合した知識の表現が可能になる。換言すると、同じオブジェクトであってもリレーションの定義の仕方を変えることによって違う知識データベースが構築される。

【 0 0 1 8 】

(4) 画像をオブジェクト値として有する複数のオブジェクトと、オブジェクト間の関係を定量的に表現した値をリレーション値として有する複数のオブジェクト間のリレーションとを蓄積したことを特徴とする知識データベース。

(5) 複数のオブジェクトと、オブジェクト間のリレーションとを蓄積した知識データベースであって、オブジェクトは画像をオブジェクト値として有する画像オブジェクトと、画像に関連付けられた自然言語情報に含まれる句をオブジェクト値として有する句オブジェクトとを含み、画像オブジェクトと句オブジェクト

との間のリレーションはそれぞれ予め定義されたリレーション値を有することを特徴とする知識データベース。

画像に関連付けられた自然言語情報に含まれる句には、当該画像を特定するための識別子（記号、符号、識別番号、ファイル名、URL等）も含まれる。

【0019】

（6）複数のオブジェクトと、オブジェクト間のリレーションとを蓄積した知識データベースであって、オブジェクトは医学分野及び／又は生物学分野で用いられる用語を含む句をオブジェクト値として有する句オブジェクトを含み、リレーションは句オブジェクトの間の関係を定量的に表現した値をリレーション値として有することを特徴とする知識データベース。

【0020】

（7）（6）記載の知識データベースにおいて、同一のオブジェクトと見なすべき複数のオブジェクトをシノニムオブジェクトグループとして登録したシノニムオブジェクト辞書を備えることを特徴とする知識データベース。

（8）（6）記載の知識データベースにおいて、オブジェクトはDNA配列をオブジェクト値として有するDNA配列オブジェクト、蛋白質配列をオブジェクト値として有する蛋白質配列オブジェクト、又は画像をオブジェクト値として有する画像オブジェクトを含み、DNA配列オブジェクト、蛋白質配列オブジェクト又は画像オブジェクトと句オブジェクトとの間のリレーションは予め定義されたリレーション値を有することを特徴とする知識データベース。

画像の例としては病変組織の顕微鏡写真を挙げることができる。

【0021】

（9）（8）記載の知識データベースにおいて、相同性が高いDNA配列をオブジェクトとして有する複数のDNA配列オブジェクト、あるいは相同性が高い蛋白質配列をオブジェクト値として有する複数の蛋白質配列オブジェクトを同一オブジェクトグループとして登録したシノニムオブジェクト辞書を備えることを特徴とする知識データベース。

【0022】

本発明の知識データベースは、知識リソースから得られるオブジェクトの言語

学的な揺らぎ、専門分野の違いによる表現の違い、実験における情報の誤差などによる情報の希釈を回避する方法として、複数のオブジェクトを同等のオブジェクトと見做し（シノニムオブジェクトグループ）、それらに関連したリレーションの合併を行う方法を用意する。シノニムオブジェクトグループに含まれる全オブジェクトは知識データベースに関連付けてシノニムオブジェクト情報として保存し、また、知識データベースにはシノニムオブジェクトグループを代表するオブジェクト（シノニム代表オブジェクト）を設定してこれを登録する。

【 0 0 2 3 】

シノニムオブジェクトグループを設定済みの知識データベースのアップデートにおいては、アップデート対象のオブジェクトがシノニムオブジェクト情報に存在する場合、対象であるシノニムオブジェクトグループのシノニム代表オブジェクトのリレーションをアップデートする。

【 0 0 2 4 】

知識データベースに対して新たにシノニムオブジェクトグループを設定する際は、シノニム代表オブジェクトを生成して知識データベースに登録し、シノニムオブジェクトグループに含まれるオブジェクトは知識データベースからは削除し、シノニムオブジェクト情報として保存する。シノニムオブジェクトグループに含まれるオブジェクトに関連していたリレーションは、全て合併してシノニム代表オブジェクトと他のオブジェクトとのリレーションとする。

【 0 0 2 5 】

句オブジェクトに対しては、指定した複数の句オブジェクトをシノニムオブジェクトグループとして登録する。シノニムオブジェクトグループに含まれる句オブジェクトの値としては、言語学的な揺らぎ、専門分野の違いによる表現の違い、頻出ミス記述などに起因する値とする。シノニム代表オブジェクトには、任意の句オブジェクトを選択する。

【 0 0 2 6 】

DNA配列オブジェクト、蛋白質配列オブジェクトにおいては、たとえば、DNA配列や蛋白質配列同士が、配列上の指定した以上の範囲において、ある値以上の相同性を持つDNA配列オブジェクト、蛋白質配列オブジェクト同士をシノニムオブ

ジェクトグループとして登録する。シノニム代表オブジェクトは、任意のDNA配列オブジェクト、蛋白質配列オブジェクトを選択するか、コンセンサス配列を作成して登録する。

複数のシノニムオブジェクトグループを登録したシノニムオブジェクト辞書を用意することにより、異なる知識データベースに対して同等シノニムオブジェクトグループを設定可能とする。

【 0 0 2 7 】

後述する知識データベースの演算においては、オブジェクト、及び、リレーションの演算に先立って、シノニムオブジェクト情報の合併とシノニムオブジェクト情報の知識データベースへの反映を行う。それにより、演算後の知識データベース内でのオブジェクトとシノニムオブジェクト情報の不整合を避けることができる。また、効果的な演算を行うために、予め、シノニムオブジェクト辞書によるシノニムオブジェクトグループの共通化を行うことが効果的である。

【 0 0 2 8 】

( 1 0 ) ( 1 ) 又は ( 6 ) 記載の知識データベースにおいて、句オブジェクトは自然言語で表現された知識（言語化知識）中の句をオブジェクト値として有し、それぞれの句が言語化知識内で近傍に存在する頻度を定量化した値を対応する句オブジェクト間のリレーションのリレーション値として有することを特徴とする知識データベース。

【 0 0 2 9 】

言語化知識中の句は、言語化知識を予め用意した 1 文字又は複数文字より成る文分離文字列により文へ分解し、続いて予め用意した 1 文字又は複数文字より成る句分離文字列により句に分解して生成してもよいし、ユーザが任意に指定したものであってもよい。また、句が言語化知識内で近傍に存在するとは、例えば文書の同じ頁内、同じ段落内、同じセンテンス内に存在することをいう。

【 0 0 3 0 】

( 1 1 ) ( 1 ) 又は ( 6 ) 記載の知識データベースにおいて、句オブジェクトは書籍の索引に含まれる句をオブジェクト値として有し、2つの句が当該書籍の同じ頁あるいは同じ段落に存在する頻度を定量化した値を対応する2つの句オブジ

ェクト間のリレーションのリレーション値として有することを特徴とする知識データベース。

【0031】

(12) (1) 又は (6) 記載の知識データベースにおいて、句オブジェクト間のリレーションは、それぞれの句が、自然言語で表現された知識の知識単位において存在するパターンの類似性（存在プロファイル）を定量化した値をリレーション値として有することを特徴とする知識データベース。

【0032】

知識単位とは、書籍等の文書の頁や文節、Webページ、DNAデータベースエントリなどを意味する。それらにおける句オブジェクトの出現パターンや出現頻度パターンによってクラスタ解析を行い、各句間の類似性距離を求め、これをリレーション値とする。

【0033】

(13) 医学分野及び／又は生物学分野で用いられる用語を含む句をオブジェクト値として有するオブジェクトと、オブジェクト相互間の関係を定量的に表現した値をリレーション値として有するオブジェクト間のリレーションとを含む知識データベースから、1 又は複数のキーワードに対して、それらをオブジェクト値として持つオブジェクト（クエリーオブジェクト）に関連するオブジェクトをオブジェクト間のリレーションを元に抽出し、クエリーオブジェクト及び抽出されたオブジェクトをその間のリレーションと共に表示することを特徴とする方法。

知識データベースは、DNA配列、蛋白質配列、蛋白質高次構造などの分子構造をオブジェクトの値として含んでいてもよい。更に、病変組織の顕微鏡写真などの画像をオブジェクトの値として含んでいてもよい。

【0034】

(14) (13) 記載の方法において、クエリーオブジェクトをオブジェクト値の入力又は検索により選択し、選択されたクエリーオブジェクト及び抽出されたオブジェクト並びにその間のリレーションをリスト表示又はグラフィカル表示することを特徴とする方法。

【0035】

(15) (13) 記載の方法において、クエリーオブジェクトに対して、強いリレーションを持つオブジェクトを階層的に求め、クエリーオブジェクト及び階層的に求めたオブジェクト並びにその間のリレーションをリスト表示又はグラフィカル表示することを特徴とする方法。

【0036】

強いリレーションを持つオブジェクトとは、オブジェクト相互間のリレーションのリレーション値が大きなオブジェクトである。どの程度のリレーション値をもって強いリレーションとするかは、ユーザが任意に指定することができる。あるいは、クエリーオブジェクトから求められたオブジェクトの数が一定数以下となるようにリレーション値が大きな方から採用する数を設定するようにしてもよい。リスト表示あるいはグラフィカル表示されたオブジェクト、あるいはリレーションにおいて、1つあるいは複数のオブジェクト、あるいはリレーションを選択可能とし、選択した1つ又は複数のオブジェクトを新たにクエリーオブジェクトとして、ターゲットオブジェクトを求めることができるようにするのが好ましい。また、選択した複数のオブジェクトをシノニムとしてシノニム辞書に登録可能としてもよい。

【0037】

(16) 自然言語により表現された医学分野及び／又は生物学分野の知識（言語化知識）より、その表現において使用されている句を抽出し、その句をオブジェクト値として有するオブジェクトを生成するとともに、各オブジェクト間の関係を定量的に表現した値をリレーション値として有するオブジェクト間のリレーションを生成し、生成したオブジェクトとオブジェクト間のリレーションとを蓄積することを特徴とする知識データベースの構築方法。

【0038】

(17) (16) 記載の知識データベースの構築方法において、言語化知識を、予め用意した1文字又は複数文字より成る文分離文字列により文へ分解し、続いて、予め用意した1文字又は複数文字より成る句分離文字列により句に分解し、これを句オブジェクトのオブジェクト値とし、2つの句が前記言語化知識内で近傍に存在する頻度を対応する2つの句オブジェクト間のリレーションのリレーシ

ョン値とすることを特徴とする知識データベースの構築方法。

【0039】

(18) (16) 又は (17) 記載の知識データベースの構築方法において、言語化知識として書籍の索引、書籍の目次、学术论文のタイトル、書籍の本文、学术论文の本文、及び／又はWebページの本文を使用することを特徴とする知識データベースの構築方法。

(19) (16) 記載の知識データベースの構築方法において、句オブジェクトのマスターデータとして書籍の索引に含まれる句を用い、2つの句が前記書籍の同じ頁に存在する頻度を対応する2つのオブジェクト間のリレーションのリレーション値とすることを特徴とする知識データベースの構築方法。

書籍としては、教科書、専門書、解説書、辞典などを利用することができる。

【0040】

(20) (16) 記載の知識データベースの構築方法において、句オブジェクトのマスターデータとして予め指定した任意の句を用い、2つの句が言語化知識内で近傍に存在する頻度を対応する2つのオブジェクト間のリレーションのリレーション値とすることを特徴とする知識データベースの構築方法。

(21) (16) 記載の知識データベースの構築方法において、言語化知識として学术论文を用い、学术论文のタイトル及び／又は本文を含む情報をネットワーク経由等により定期的を取得し、取得した情報から句オブジェクト及びリレーションを抽出し、知識データベースをアップデートすることを特徴とする知識データベースの構築方法。

【0041】

(22) (16) 記載の知識データベースの構築方法において、同一のオブジェクトと見なす複数のオブジェクト及びそれらに関連したリレーションを併合することを特徴とする知識データベースの構築方法。

同一のオブジェクトと見なすオブジェクトはユーザが指定することができる。

(23) DNA配列を含む情報からDNA配列を抽出し当該DNA配列をオブジェクト値とするDNA配列オブジェクトを生成し、2つのDNA配列オブジェクト間に、対応する2つのDNA配列間の関係を相同性計算あるいは実験的な測定により求められた

定量的な値をリレーション値として有するリレーションを生成し、生成したDNA配列オブジェクト及びDNA配列オブジェクト間のリレーションを蓄積することを特徴とする知識データベースの構築方法。

【0042】

(24) (23) 記載の知識データベースの構築方法において、DNA配列に関連付けられた自然言語情報に含まれる句を抽出しそれをオブジェクト値とする句オブジェクトを生成し、句オブジェクトと対応するDNA配列オブジェクトの間に定義されたりレーション値を有するリレーションを設定することを特徴とする知識データベースの構築方法。

【0043】

DNA配列に加えて蛋白質配列も同様にオブジェクト化してもよい。オブジェクト及びリレーションを抽出して知識データベースを構築するのに用いる言語化知識としては、DNA配列、蛋白質配列又はその両方とそれらの付加情報を含む言語化知識、DNA配列、蛋白質配列又はその両方とそれらの付加情報を蓄積したデータベースを使用することができる。また、HTML、XMLなどの構造化記述表現で表現された知識を使用してもよい。

【0044】

(25) (24) 記載の知識データベースの構築方法において、DNA配列とそれに関係付けられた自然言語情報を含む情報をネットワーク経由等により定期的に取り得し、取得した情報からDNA配列オブジェクト、句オブジェクト及びリレーションを抽出し、知識データベースをアップデートすることを特徴とする知識データベースの構築方法。

【0045】

(26) (23) 記載の知識データベースの構築方法において、相同性が高いDNA配列をオブジェクト値として有する複数のDNA配列オブジェクトを同一オブジェクトと見なし、同一オブジェクトと見なした複数のDNA配列オブジェクト及びそれらに関連したリレーションを併合することを特徴とする知識データベースの構築方法。

(27) 自然言語により表現された句をオブジェクト値として有する複数のオブ

ジェクトと、オブジェクト間の関係を定量的に表現した値をリレーション値として有する複数のオブジェクト間のリレーションとを蓄積した第1及び第2の知識データベース間で演算を行い、第3の知識データベースを生成することを特徴とする知識データベースの生成方法。

知識データベース間の演算としては、和、差などがある。

【0046】

(28) 自然言語により表現された句をオブジェクト値として有する複数のオブジェクトと、オブジェクト間の関係を定量的に表現した値をリレーション値として有する複数のオブジェクト間のリレーションとを蓄積した第1及び第2の知識データベースを用い、第1及び第2の知識データベースにおける注目するリレーションのリレーション値の違いを比較するグラフを表示することを特徴とする方法。

時系列に意味のある一連の知識リソースより生成された複数の知識データベースを対象として、この方法を実行すると、知識リソースに含まれる知識の時系列の変化を比較するグラフを表示することができる。

【0047】

(29) (27) 又は (28) 記載の方法において、複数の知識データベース間で、各知識データベースに含まれる句オブジェクト間のリレーションのリレーション値を標準化することを特徴とする方法。

標準化の方法としては、知識データベースの元データである知識リソースにおけるオブジェクトの出現頻度によって標準化する方法、知識データベース中の注目するリレーションの値が同じ値になるように標準化する方法などがある。知識データベース間の演算や、知識データベースの比較においては、予め、対象の知識データベースにおけるシノニムオブジェクトグループを統一することで、演算後又は比較における知識データベース内でのオブジェクトシノニム情報の不整合を防ぐことが望ましい。

【0048】

本発明によると、学術論文などの言語化された知識や、DNA配列、蛋白質配列の様に記号化された知識を、句、事象、記号といったオブジェクトと、それらの

間のリレーションにより情報処理可能な知識骨格構造として再構築し、それを利用してある知識に関連する知識を獲得することにより、知識の説明、知識の表現、知識の発見が可能となる。

【0049】

【発明の実施の形態】

以下、図面を参照して本発明の実施の形態を説明する。

図1は、本発明のシステム構成図である。本発明のシステムは、教科書、専門書など印刷された医学・生物学の言語化知識データ100、学術論文や知識情報など、ネットワークを介してアクセス可能なリモートシステム上で公開され、順次更新される言語化知識データ101、DNA配列情報や蛋白質配列情報およびそれに関連付けられた言語化知識など、ネットワーク109を介してアクセス可能なリモートシステム上で公開され、順次更新される学術情報データベース102、印刷された言語化知識をイメージとして読み取るスキャナ装置103、スキャナ装置を介して得たイメージをテキストに変換し、変換した言語化知識の他、ネットワークを介して得た言語化知識やDNA配列や蛋白質配列をリソースとして知識データベースを構築し、知識データベースを利用する処理を行う処理部104、処理部により構築された知識データベースを蓄積するデータベース105、医学・生物学知識の知識要素同士の関連を視覚化して表示するための表示装置106、本システムへの値の入力や選択の操作を行うためのキーボード107やマウス108などの入力装置から構成される。

【0050】

図2は、最も単純な知識データベースの概念図である。すなわち、本発明における知識データベース203は、オブジェクト201とその間のリレーション202によって構成される。リレーションは、知識データベース内で定義された2つのオブジェクト間の方向性を持たない関係値（スカラー値）である。また、図3の知識データベースの概念図に示す様に、リレーション202は、知識データベース301に存在する全てのオブジェクト201間に存在するものとする。

【0051】

2つの知識データベースを合併（知識データベースの和）することにより、知

知識データベースの情報量を増加したり、異なる専門分野にまたがって存在する潜在的な知識を含む知識データベースを生成したりすることが可能である。また、2つの知識データベース間で差を取る（知識データベースの差）ことにより、知識データベース間でユニークな知識や知識の偏りを含む知識データベースを生成することができる。こうした知識データベース間の処理（知識データベースの演算）は、演算元の両知識データベースの持つ全オブジェクトを持ち、また、演算元の両知識データベース由来のリレーションを元に算出したリレーション値又はオブジェクトの値より算出したリレーション値を持つ知識データベースを生成することで実現できる。知識データベースの演算におけるリレーション値の算出においては、演算元の知識データベースに存在しなかったリレーションについてはリレーションを値0として算出を行う。また、演算元の知識データベースにおいて「対象となる2つのオブジェクトを同一とみなす」とする値を持つリレーションについては、オプションにより、リレーションの算出結果としてこの値を維持するか、又は異なる値に変換後リレーションの算出を行う。

#### 【0052】

知識データベースの和においては、リレーション値の算出は、演算元の2つの知識データベースにおけるリレーション値の和により行えばよい。

知識データベースの差においては、リレーション値の算出は、演算元の2つの知識データベースにおけるリレーション値の差により行えばよい。算出において、演算の左辺のリレーション値が右辺のリレーション値よりも小さい場合は、リレーション値の算出結果として0を得ることとすればよい。

#### 【0053】

図4は、知識データベースの合併の概念図である。2つの知識データベースを合併する場合、合併対象の両知識データベースにおいてオブジェクトの比較を行い、同一であるとみなされたオブジェクト同士を合併する。その際、合併する2つのオブジェクト間のリレーションは削除し、合併する2つのオブジェクトと他の全てのオブジェクトとの間のリレーション値は、合併するオブジェクトの種類に応じて指定した関数又は漸化式により算出する。

#### 【0054】

たとえば、知識データベースAにおけるオブジェクトaとbの間のリレーションをAabと表現し、また、知識データベースAと知識データベースBの合併で生成される知識データベースABにおけるオブジェクトa, b間のリレーションを算出する関数を $ABab = f(Aab, Bab)$ とすると、知識データベースA (4 0 1) と知識データベースB (4 0 2) の合併によって生成される知識データベースAB (4 0 3) において、合併対象の両知識データベースに共通して存在していたオブジェクト4 0 6, 4 0 7間のリレーションであるABcd (4 0 9) は、 $ABcd = f(Acd, Bcd)$ として、また、合併対象の知識データベースの一方にのみ存在していたオブジェクト間のリレーション、たとえばABac (4 1 0) は、 $ABac = f(Aac, 0)$ として、また、合併により新たに発生したリレーションであるABae (4 1 1) は、 $ABae = f(0, 0)$ として算出し、リレーション値とすることができる。

#### 【 0 0 5 5 】

このとき、たとえば、句オブジェクトと他の種類のオブジェクトとの間のリレーションは $f(Aab, Bab) = Aab + Bab$ により値の算出を行ったり、何れか一方のリレーション値を維持したりする。また、DNA配列オブジェクト、蛋白質配列オブジェクト同士のリレーションにおいては、合併後のオブジェクトの値によってリレーションの値を再計算したり、何れか一方の知識データベース由来のリレーション値を維持したりする。また、画像オブジェクト同士のリレーションにおいては、合併後のオブジェクトの値によってリレーションの値を再計算したり、何れか一方の知識データベース由来のリレーション値を維持したりする。知識データベースの合併において、オブジェクトが同一であることの判定方法は、句オブジェクトにおいては、たとえば、句の一致するものを同一とみなす。また、DNA配列オブジェクト、蛋白質配列オブジェクトにおいては、たとえば、相同性計算を行い、閾値として相同性スコアや最大一致並置における一致範囲長と一致率などを設定し、閾値を超えるオブジェクトの組を同一とみなす。また、画像オブジェクトにおいては、たとえば、パターンマッチングを行い、閾値として設定した値を超える画像オブジェクトの組を同一とみなす。合併後のオブジェクトの値は、句オブジェクトにおいては変わらず、また、DNA配列オブジェクト、蛋白質配列オブジェクトにおいては、たとえば、いずれか一方の値を合併後の値としたり、合併

対象の知識データベースにおける2つのオブジェクトの値よりコンセンサス配列を作成しそれを新しい値としたりすればよい。また、画像オブジェクトにおいては、たとえば、いずれか一方の値を合併後の値とすればよい。

## 【0056】

図5は、知識データベース内におけるオブジェクトの合併の概念図である。知識データベース内の2つのオブジェクトを合併する場合、合併対象の両オブジェクト間のリレーションを削除し、両オブジェクトと他の全てのオブジェクトとの間にそれぞれ存在したリレーションは合併し、再計算する。

## 【0057】

知識データベースCにおけるオブジェクトaとb間のリレーションを $C_{ab}$ と表現し、また、オブジェクトaとオブジェクトbの合併で生成されるオブジェクトabとオブジェクトcとの間のリレーションを算出する関数を $C_{ab,c} = f(C_{ac}, C_{bc})$ とすると、図5における処理前の知識データベースC(501)のオブジェクトbとeを合併した結果生成する、合併後の知識データベースC(502)に存在するオブジェクトbeとオブジェクトcとの間のリレーションである $C_{be,c}$ は、 $C_{be,c} = f(C_{bc}, C_{ce})$ として算出しリレーション値とする。たとえば、句オブジェクトと他の種類のオブジェクトにおけるリレーションは $f(C_{bc}, C_{ce}) = C_{bc} + C_{ce}$ により値の算出を行う。また、たとえばDNA配列オブジェクト、蛋白質配列オブジェクト同士のリレーションにおいては、合併後のオブジェクトの値によって、リレーションの値を再計算する。合併後のオブジェクトの値は、句オブジェクトにおいては変わらず、また、DNA配列オブジェクト、蛋白質配列オブジェクトにおいては、たとえば、いずれか一方の値を合併後の値としたり、合併対象の知識データベースにおける2つのオブジェクトの値よりコンセンサス配列を作成しそれを新しい値としたりすればよい。

## 【0058】

図6は、種類の異なるオブジェクトを含む知識データベースの例を示す図である。ここでは、リレーションの値として、対象となる2つのオブジェクトを同一とみなす場合は-1、オブジェクト間に関係が無い場合は0、オブジェクト間に関係がある場合は0より大きな値とし、値が大きいほどオブジェクト間の関係が

高いものとしている。ただし、図を見易くするため、種類の異なるオブジェクト間のリレーションについては、値が0の物の素描を省略している（たとえば、句オブジェクト602とDNA配列オブジェクト607の間のリレーションなど）。

【0059】

知識データベース601には、句オブジェクト602～606、及び、DNA配列オブジェクト607～609が含まれ、それぞれの種類のオブジェクト内でリレーションを持つほか、異なる種類のオブジェクト間でもリレーションを持つ。たとえば、オブジェクト604と607、オブジェクト605と608、オブジェクト606と609の間のリレーション612, 613, 614は、値が-1であり、これらのオブジェクトは同一であるとみなすことを示している。

【0060】

図7は、図6の例において示した知識データベースより、リレーション値が10以上であるリレーションにより関連付けられた句オブジェクトを含む知識データベース701を抽出した例である。図6においては、句オブジェクトに関してだけで評価すると、オブジェクト604と606のリレーション610の値は0であり、関連が無いことを示しているが、これを、オブジェクト604, 606について、それぞれリレーション612, 614により同一とみなすと定義されたオブジェクト607, 609の間のリレーション611と合併することにより、値20を持つリレーション702とすることができる。このように、構築した知識データベースからリレーション値をもとに、異なる種類のオブジェクトを介した知識情報を得ることが可能である。

【0061】

言語化知識である知識リソースからの句オブジェクトの抽出は、たとえば、言語化知識を予め用意した1文字又は複数文字より成る文分離文字列により文に分解し、続いて、予め用意した1文字又は複数文字より成る句分離文字列により句に分解し、これを句オブジェクトとする方法により行うことができる。その際のリレーションの設定は、たとえば、句同士が同じ頁又は文に存在する頻度によって値を設定すればよい。

【0062】

図8は、言語化知識を知識のリソースとした場合において、そこからの句オブジェクトの切り出し、リレーションの設定を行う手順の一例をフローチャートにして示したものである。

処理の開始(801)後、まず、作成対象の知識データベースを用意して初期化を行う(802)。この時点では、知識データベース中には、オブジェクトも、よってリレーションも存在しない。続いて、知識リソースよりデータ単位として頁を取り込む(803)。本例ではデータ単位として、頁と文を考え、2つのオブジェクトが同じ頁に存在した場合、そのリレーションのスコアをP0、同じ文に存在した場合、リレーションのスコアをP1としている。頁が存在するか判定し(804)、頁が存在する場合は、その頁の情報を作成対象の知識データベースへ加える処理を行う。その手順としては、まず、頁の知識を蓄えるための知識データベース(頁用知識データベース)を用意して初期化する(805)。

#### 【0063】

続いて、各頁より文を切り出す(806)。文の切り出しには、予め登録した、1文字又は数文字からなる「文切り出し文字列」を用いる。文切り出し文字列には、例えば、「.(ピリオド)」、「(改行)」を登録する。文切り出し文字列にて頁に含まれる言語化知識を分断してそれぞれを文とする。文が存在するか判定し(807)、文が存在する場合、文の知識を蓄えるための知識データベース(文用知識データベース)を用意して初期化する(808)。続いて、各文より句を切り出し、句オブジェクトを生成する(809)。句の切り出しには、予め登録した、1文字又は数文字からなる「句切り出し文字列」を用いる。句切り出し文字列には、例えば、「and」、「that」、「(Tab)」を登録する。句切り出し文字列にて文に含まれる言語化知識を分断し、分断した文字列の先頭と末尾よりスペースを削除してそれぞれを句とする。句オブジェクトが存在するか判定し(810)、存在する場合は文用知識データベースへ登録(811)し、次の句オブジェクトの切り出しを行う。句オブジェクトの文用知識データベースへの登録においては、リレーション値を0とする。句オブジェクトが存在しない場合、対象文のオブジェクト抽出が全て完了したとして、文用知識データベースの全リレーションに値P1を加算(812)した後、頁用知識データベースに合併(8

13) し、次の文の切り出しを行う。

【0064】

文が存在するか判定し(807)、文が存在しない場合は、対象頁のオブジェクト抽出が全て完了したとして、頁用知識データベースの全リレーションに値P0を加算(814)した後、作成対象の知識データベースに合併(815)し、次の頁の取り込みを行う。頁が存在するか判定し(804)、存在しない場合は処理を終了し、作成対象の知識データベースの作成を完了する(816)。以上の手順によって、言語化知識からの句オブジェクトの切り出し、及び、リレーションの設定を行うことができる。

【0065】

図9は、言語化知識からの句オブジェクトの切り出しと、リレーションの設定の方法の他の例として、論文のタイトルを知識のリソースとした場合を示した図である。論文のタイトル901は、インターネットを通じた情報の閲覧、検索や、情報配信サービス、また、書籍などから入手可能である。そうした論文のタイトルを切り出し(902)、句切り出し文字列904によりタイトルを句903に分離し、各句をオブジェクト905として、各句の間のリレーション906を設定する。ここでは、リレーション値1を設定している。

【0066】

言語化知識である知識リソースからの句オブジェクト抽出の他の方法として、たとえば、予め教科書、専門書、解説書、辞典等の書籍の索引に掲載されている句(用語、熟語、文節など)を集め、これを句オブジェクトのマスターデータとして用い、対象知識リソースから同じ句を抽出する方法がある。その際のリレーションの設定は、例えば、句と同時に索引から抽出した頁番号を元に、句同士が同じ頁に存在する頻度によって値を設定すればよい。

【0067】

図10は、書籍の索引を知識のリソースとした場合、そこから、句オブジェクトの切り出しとリレーションの作成を行う手順の一例をフローチャートにして示したものである。

処理の開始(1001)後、まず、作成対象の知識データベースを用意して初

期化を行う(1002)。この時点では、知識データベース中には、オブジェクトも、よってリレーションも存在しない。次に、出現パターンデータの初期化を行う(1003)。出現パターンデータには、索引に含まれる各句がどのような頁番号を持つかを格納する。続いて、知識リソースより索引を含む頁を取り込む(1004)。頁が存在するか判定し(1005)、頁が存在する場合は、索引に含まれる句を切り出す(1006)。句が存在するか判定し(1007)、句が存在する場合は出現パターンデータに登録する(1008)。句に連なる頁番号を切り出し(1009)、頁番号が存在するか判定し(1010)、頁番号が存在する場合は出現パターンデータに登録(1011)し、次の頁番号を切り出す(1009)。頁番号が存在するか判定し(1010)、頁番号が存在しない場合は次の句の切り出し(1006)を行う。

#### 【0068】

句が存在するか判定し(1007)、句が存在しない場合は次の頁の取り込み(1004)を行う。頁が存在するか判定し(1005)、頁が存在しない場合は、全ての索引の読み取りが完了したと判断し、各句を句オブジェクトとして知識データベースに登録(1012)し、また、出現パターンデータを参照して、各句の組が同じ頁に存在する数をリレーション値として登録(1013)して、作成対象の知識データベースの作成を完了する(1014)。

#### 【0069】

図11は、索引を知識のリソースとして、そこから、句オブジェクトを切り出し、また、リレーションを設定する方法を図示している。まず、知識のリソースである索引1101より、句の部分1102及び頁番号の部分1103をそれぞれ抽出し、句を切り出し、また、頁出現パターンからなる出現パターンデータ(1104)を生成する(図10における1001-1011の処理)。出現パターンデータをもとに、句を句オブジェクトとし、また、2つの句がそれぞれ同じ頁に出現する回数をカウントしてリレーション値とする(1105)(図10における1012-1013の処理)。これを図示すると1106となる。

また、言語化知識である知識リソースからの句オブジェクト抽出の他の方法として、たとえば、予め指定した任意の句(用語、熟語、文節など)を句オブジェク

トのマスターデータとして用い、対象知識リソースから同じ句を抽出する方法がある。この場合は、まず、指定した句により、知識単位における句の存在を示す索引を作成し、前述した書籍の索引を知識リソースとした場合と同じ方法により、句オブジェクトの切り出しとリレーションの設定を行うことができる。

また、知識単位における出現パターンよりリレーション値を得る方法として、上記2例に示した各知識単位における句オブジェクトの存在の有無ではなく、存在量を集計し、各知識単位における存在プロファイルとしてこれをクラスタ解析することにより、各知識単位における句オブジェクトの存在プロファイルの類似性を定量化し、これを各オブジェクト間のリレーション値とすることも有用である。この方法によって、知識単位における出現パターンよりリレーション値を得る方法において、特に、学術論文、Webページなど、比較的情報量の多い知識単位を対象とした場合において、情報の希釈を回避することができる。

#### 【0070】

句オブジェクトにより構成される知識データベースにおいては、知識リソースである言語化知識の選択によって、知識の情報密度が大きく左右される。すなわち、言語化知識に含まれる冗長な情報、雑多な情報は、知識データベースにおけるノイズ情報となり、知識データベースから意味情報を抽出する際の妨げとなる。知識データベース構築に使用する知識リソースを、目的に応じて選択することで、ノイズ情報の少ない知識データベースを構成することが可能となる。

#### 【0071】

すなわち、言語化知識の知識リソースとして、教科書、専門書、解説書などの書籍の索引を対象とした場合、書籍の著者や編者により厳選された句のみを知識データベースに取り込むことが可能となる。一般の言語化知識においてはノイズとなる“DNA”、“遺伝子”などの頻出語も、索引においては単独の句としては採用されず、知識を構成する句としてのみ載っているため、句オブジェクトとしての利用が可能となる。また、言語化知識の知識リソースとして、教科書、専門書、解説書などの書籍の目次を対象とした場合、執筆時点での知見に基づいた事実のみを述べているため、厳選された句と共に、選択的に強いリレーションを得ることができる。

【0072】

また、言語化知識の知識リソースとして、学术论文のタイトルを対象とした場合、医学・生物学における学术论文の特徴として、タイトルにはその論文が証明したり、説明したりする知識の最も主要な要素が含まれるため、厳選された句と共に、選択的に強いリレーションを得ることができる。学术论文のタイトルは、コンピュータネットワークを介して取得することが可能であり、定期的に最新の学术论文を取得し、タイトルより自動処理により句オブジェクトとリレーションを抽出し、知識データベースをアップデートする仕組みにより、最新の知識を反映した知識データベースを維持することが可能である。

【0073】

また、言語化知識の知識リソースとして、教科書、学术论文等の本文を対象とすることにより、均一な句オブジェクトを得ることが可能となる。すなわち、こうしたリソースは、一般の言語化知識において情報処理の妨げとなる同意語や同義語（シノニム）が少なく、それらによる知識情報の希釈を避けることができる。たとえば、“マイクロアレイ”、“DNAチップ”、“バイオチップ”といった用語は、いずれも同一の事象を指す場合があるが、一般的にこうしたリソースにおいては用語が統一されている。学术论文はコンピュータネットワークを介して取得することが可能であり、定期的に最新の学术论文を取得し、本文より自動処理により句オブジェクトとリレーションを抽出し、知識データベースをアップデートする仕組みにより、最新の知識を反映した知識データベースを維持することが可能である。

【0074】

DNA配列オブジェクトや蛋白質配列オブジェクトは、DNA配列や蛋白質配列と共に言語化知識を含む知識リソースより抽出する。通常、DNA配列あるいは蛋白質配列は、そのDNA配列あるいは蛋白質配列を特定する識別子、及び、DNA配列あるいは蛋白質配列の由来や機能を記述した言語化知識と共にデータ化されている。そうした知識リソースより、DNA配列あるいは蛋白質配列をDNA配列オブジェクトあるいは蛋白質配列オブジェクトとして抽出し、相互の間にリレーションを設定する。また、それらの配列に付随する識別子を句オブジェクトとして抽出し、対

応するDNA配列オブジェクトあるいは蛋白質配列オブジェクトとの間のリレーションに「対象となる2つのオブジェクトを同一とみなす」とする値を設定する。また、言語化知識から抽出したその他の句オブジェクトとDNA配列の識別子をオブジェクト値とする句オブジェクトや蛋白質配列の識別子をオブジェクト値とする句オブジェクトとの間、及び、言語化知識から抽出したその他の句オブジェクト相互の間にリレーションを設定する。これにより、DNA配列オブジェクト、蛋白質配列オブジェクトと句オブジェクトを共に含む知識データベースが構築される。この知識データベースは、句オブジェクトを含むため、他の句オブジェクトを含む知識データベースと容易に合併可能である。

#### 【 0 0 7 5 】

DNA配列情報や蛋白質配列情報を蓄積したデータベースを知識リソースとする場合は、DNA配列エントリー単位あるいは蛋白質配列エントリー単位に、上記の操作を行い、エントリー単位に生成した知識データベースを全て合併してデータベースの知識データベースとする。DNA配列あるいは蛋白質配列、又はその両方と言語化知識を共に含む知識リソースは、コンピュータネットワークを介して取得することが可能であり、定期的に最新の情報を取得し、DNA配列オブジェクト、蛋白質配列オブジェクト、句オブジェクト及びリレーションを抽出し、定期的に取得した知識リソースに対する知識データベースを作成し、マスターの知識データベースへ合併することで、最新の知識を反映した知識データベースを維持することが可能である。

画像オブジェクトにより構成される知識データベースにおいては、たとえば画像オブジェクト間の特性値解析の結果や、画像取得状況をリレーション値とする。例えば、時系列に取得した画像において、取得の次期の近い画像間ほど高いリレーション値とする。目的に応じて画像オブジェクト間のリレーション値を用意した知識データベースを、他の画像オブジェクトを含む知識データベースへ合併することによって、画像オブジェクト間に特徴あるリレーション値を提供することが可能である。例えば、生体組織の成長や老化、病変における時系列変化を捉えた一連の画像を画像オブジェクトとして持つ知識データベースを用意し、複数の検体よりサンプルとして取得した生体組織画像を含む知識データベースと合併

することにより、各検体間の生体組織画像からみた定量的関係を含む知識データベースを構築することが可能である。

【 0 0 7 6 】

画像オブジェクトは、画像と画像に関連付けられた言語化知識とを含む知識リソースより抽出する。例えば、画像を特定する識別子を持つ画像ファイルより、画像を画像オブジェクトとして、識別子を句オブジェクトとして抽出する。また、識別子により結び付けられて画像の内容を記述した言語化知識より、句オブジェクトを抽出する。そして、画像オブジェクトとその画像の識別子をオブジェクト値とする句オブジェクトとの間のリレーションに「対象となる2つのオブジェクトを同一とみなす」とする値を設定する。また、識別子により結び付けられて画像の内容を記述した言語化知識から抽出した句オブジェクト相互の間、及び、当該句オブジェクトと識別子をオブジェクト値とする句オブジェクトとの間にリレーションを設定する。これらのオブジェクト及びリレーションを含む知識データベースは、句オブジェクトを含むため、他の句オブジェクトを含む知識データベースと容易に合併可能である。

こうした、句オブジェクトと画像オブジェクトを共に含む知識データベースは、Webページを知識リソースとすることによっても生成される。Webページを識別するURLや、Webページに含まれる画像を識別するURL、Webページに含まれる言語化知識等を句オブジェクトとし、Webページ内の画像（画像オブジェクト）との間にリレーションを設定することで、他の句オブジェクトを含む知識データベースと容易に合併可能である。

【 0 0 7 7 】

知識リソースとして、HTML、XMLなどの構造化記述言語で表現された知識も利用可能とする。その場合、知識リソースの表現形式とその定義に応じて、どういった情報をどの種類のオブジェクトとして抽出し、リレーションを設定するかを用意し、それに従って利用する。

【 0 0 7 8 】

図 1 2 は、異なる種類のオブジェクトが存在する知識データベースを生成する例として、DNAデータベースエントリを知識のリソースとして、そこから、DNA配

列オブジェクトとそれに関連した句オブジェクトを切り出し、リレーションを設定する方法の例を示した図である。

#### 【 0 0 7 9 】

典型的なDNAデータベースエントリ 1 2 0 1 には、そのDNAが何物であることを説明した部分 1 2 0 2、そのDNAを特定する識別子 1 2 0 3、DNA配列 1 2 0 4 が含まれる。それぞれの部分をDNAデータベースエントリより切り出した後、DNAが何物であることを説明した部分 1 2 0 2 は、論文のタイトル同様、句分離文字列 1 2 0 6 によって、句 1 2 0 5 に分解し、句オブジェクト 1 2 0 7 とその間のリレーションを生成する。ここでは、リレーション値を 1 としている。そのDNAを特定する識別子 1 2 0 3 は、そのまま句オブジェクトとして、同じDNAデータベースエントリのDNAが何物であることを説明した部分から生成した全ての句オブジェクトとの間にリレーションを設定する。ここでは、リレーション値を N としている。最後に、DNA配列 1 2 0 4 は、そのままDNA配列オブジェクト 1 2 1 0 とし、そのDNAを特定する識別子 1 2 0 3 から生成した句オブジェクトとの間に、2つのオブジェクトを同一とみなすリレーション 1 2 1 1 を設定する。ここでは、同一とみなす値として -1 を設定している。

#### 【 0 0 8 0 】

図 1 3 は、異なる種類のオブジェクトが存在する知識データベースを生成する例として、画像ファイルと画像を説明した言語化知識を知識のリソースとして、そこから、画像オブジェクトとそれに関連した句オブジェクトを切り出し、リレーションを設定する方法の例を示した図である。

#### 【 0 0 8 1 】

画像ファイル 1 3 2 1 には、画像 1 3 2 2 と画像の識別子としてのファイル名 1 3 2 3 が含まれる。また、画像の説明 1 3 2 4 には、画像の識別子としての画像ファイルのファイル名 1 3 2 5 と画像を説明する言語化知識 1 3 2 6 が存在する。それぞれの部分を切り出した後、画像を説明する言語化知識は、論文のタイトル同様、句分離文字列 1 3 2 8 によって、句 1 3 2 7 に分解し、句オブジェクト 1 3 2 9 とその間のリレーションを生成する。ここでは、リレーション値を 1 としている。ファイル名は、画像を特定する識別子として、句オブジェクト 1 3

31とし、画像を説明する言語化知識から生成した全ての句オブジェクトとの間にリレーション1330を設定する。ここではリレーション値をNとしている。画像は、画像配列オブジェクト1332とし、画像を特定する識別子から生成した句オブジェクト1331との間に、2つのオブジェクトを同一とみなすリレーション1333を設定する。ここでは、同一とみなす値として-1を設定している。

【0082】

図14は、Webページから知識データベースを生成する例の図である。Webページ1441には、WebページのURL1442、Webページに含まれる言語化知識1443、画像1444とそのURL1445などが含まれる。それぞれの部分を切り出した後、WebページのURLは句オブジェクト1446とし、Webページに含まれる言語化知識は、論文のタイトル同様、句分離文字列1448によって、句1447に分解し、句オブジェクト1449とその間のリレーションを生成する。ここでは、リレーション値を1としている。画像は、画像オブジェクト1451とし、画像のURLは、画像を特定する識別子として、句オブジェクト1452とし、その間に、2つのオブジェクトを同一とみなすリレーション1453を設定する。ここでは、同一とみなす値として-1を設定している。Webページに含まれる言語化知識から生成した全ての句オブジェクト及び画像のURLから生成した句オブジェクトと、WebページのURLより生成した句オブジェクトとの間にリレーション1450を設定する。ここではリレーション値をNとしている。

【0083】

図15は、知識データベースにおけるデータの格納方式、及び、知識データベースに関連した情報として、シノニムオブジェクトの情報、及び、シノニムオブジェクト辞書を格納する方式を、リレーショナルデータベースにおける場合を例として示した図である。

【0084】

本データ格納方式においては、複数の知識データベースを格納可能としている。すなわち、一つの知識データベース1506は、知識データベーステーブル1501の1レコードによって他と識別される。各知識データベースにはユニーク

なIDとして知識データベース識別子 (ndb#id) が付与され、オブジェクトテーブルの各レコードは、知識データベース識別子を持つことによって、どの知識データベースに所属するかが識別可能となる。オブジェクトレコードには、知識データベース識別子の他に、オブジェクトタイプ (obj#type)、オブジェクト識別子 (obj#id)、オブジェクトの値 (obj#value)、オブジェクトカウント (obj#count) を格納可能とする。オブジェクト識別子は、知識データベース内でユニークな値とし、また、異なる知識データベースにおいても、同じオブジェクト識別子を持つオブジェクトは同じ値を持つものとする。オブジェクトタイプには、句オブジェクトやDNA配列オブジェクトといったオブジェクトの種類を示す情報を格納し、オブジェクトの値には、オブジェクトの値をテキストやバイナリで格納する。また、オブジェクトカウントには、句オブジェクトにおけるリレーション値の標準化に使用するために、知識リソースにおける対象オブジェクトの出現頻度を記録する。

## 【 0 0 8 5 】

リレーションテーブル 1 5 0 3 には 2 つのオブジェクト識別子を格納可能とし、どのオブジェクト間のリレーションであるかを識別可能とする。また、リレーション値 (rel#value) を格納可能とする。シノニムオブジェクト情報テーブルには、シノニムを持つオブジェクトを指し示すオブジェクト識別子と、シノニムの値を格納する。シノニムの値はオブジェクトと同等の値であるため、オブジェクトの値と同じフィールド名としてある。

## 【 0 0 8 6 】

また、シノニムオブジェクト辞書テーブル 1 5 0 5 を用意している。シノニムオブジェクト辞書テーブル 1 5 0 5 には、シノニムのグループを識別するためのオブジェクト識別子、シノニムの値、シノニムの付加情報 (syn#attr) を格納する。シノニムの識別子及びシノニムの値はオブジェクトと同等の情報であるため、オブジェクト識別子及びオブジェクトの値と同じフィールド名としてある。シノニム付加情報には、たとえば、対象のレコードにおけるシノニムがシノニムのグループの内での代表シノニムであるかどうかの情報を格納する。

## 【 0 0 8 7 】

図 1 6 は、知識データベースに格納する情報の例を示す図である。オブジェクトの情報 1 6 0 2 としては、オブジェクトの番号、種類、値、出現回数が、また、リレーション 1 6 0 3 としては、オブジェクトの番号によってリレーション値のマトリックスが、また、これらを含む知識データベース 1 6 0 1 の名前として「Sample NDB」が与えられていることを示している。

## 【 0 0 8 8 】

図 1 7 は、図 1 6 において示した情報を、図 1 5 で示した方式で格納する例を示す図である。知識データベーステーブル 1 7 0 1 には、知識データベース識別子と知識データベース名を持った、対象の知識データベースを示す 1 つのレコードを、オブジェクトテーブル 1 7 0 2 には、所属する知識データベースを示す知識データベース識別子、オブジェクト識別子、オブジェクトタイプ、オブジェクトの値とオブジェクトカウントを、リレーションテーブル 1 7 0 3 には、それぞれ、どのオブジェクト間のリレーションであるかを示す 2 つのオブジェクト識別子とリレーションの値を格納している。ここでは、値が 0 のリレーションはリレーションテーブルには格納しないこととしている。

## 【 0 0 8 9 】

複数の知識データベースより、同じ値を持つ複数の句オブジェクトを選出し、それらの句オブジェクト間のリレーション値を抽出し、これを比較することにより、複数の知識データベースにおける知識の構成の違いを知ることができる。ここで対象とする句オブジェクトは、利用者が任意の興味ある句オブジェクトを選択するか、あるいは、対象となる複数の知識データベースに含まれる任意の値以上のリレーション値を持つリレーションに関わる句オブジェクトを選択するか、あるいは、その 2 つの方式を組み合わせることで選択する。比較を行う方法としては、たとえば、2 つの句オブジェクトの組によって明示される対象リレーションを一方の軸に、リレーション値を他方の軸に取って、知識データベース毎に色、線種、マーカなどの異なる折れ線を引く二次元グラフを書く。グラフの表示においては、ある知識データベースにおける対象リレーションのリレーション値によって、対象リレーションの軸をソートすることで、他の知識データベースの値との比較を容易にすることができる。

## 【 0 0 9 0 】

また、時系列に意味のある知識データベース、たとえば、年毎、月毎など、一定期間毎に集められた情報を知識リソースとして生成した知識データベースなどを対象とした知識データベースの比較においては、時系列でソートした知識データベースを一方の軸に、リレーション値を他方の軸に取って、2つの句オブジェクトの組によって明示される対象リレーション毎に色、線種、マーカなどの異なる折れ線を引く二次元グラフを書く。こうしたグラフの素描により、知識リソースにおける知識の時系列の変化を容易に観察することが可能となる。

## 【 0 0 9 1 】

図 1 8 は、複数の知識データベース間における、あるオブジェクトに関連したリレーションの値の変化を示すグラフの例を示す図である。

図では、6つの知識データベース (ndb#95, ndb#96, ndb#97, ndb#98, ndb#99, ndb#00) において、「遺伝子A」の値を持つオブジェクト 1 8 0 5 に対して、いずれかの知識データベースにおいて強いリレーションを持つオブジェクト (ターゲットオブジェクト) を抽出し、グラフ表示したものである。縦軸にリレーション値 1 8 0 2、横軸に知識データベース 1 8 0 3 を取り、いずれかの知識データベースにおいて「遺伝子A」の値を持つオブジェクトに対して 0 より大きなリレーション値を持っていた6つのオブジェクト 1 8 0 6 について、それぞれ、値を線グラフ 1 8 0 4 で表している。この方法により、注目したオブジェクトに対する他のオブジェクトの関係が、異なる知識データベースにおいて、如何に異なるかを知ることができる。また、対象となる知識データベースが時系列に変化する知識リソースから作成されたものである場合、横軸の知識データベースを元となった知識リソースに応じて時系列に並べることで、注目したオブジェクトに関連したオブジェクトについて、時系列の関係の変化を見ることができる。

## 【 0 0 9 2 】

異なる知識リソースから生成した、句オブジェクトを含む知識データベースの演算を行う際、あるいは、知識データベースの比較など、知識データベース間の情報の差に着目した知識データベースの利用を行う際は、演算もしくは利用対象となる知識データベース間での句オブジェクト間のリレーションの影響力や情報

量を一定にするために、演算又は利用前に、演算もしくは利用対象の知識データベース間でリレーション値の標準化を行う。すなわち、各知識データベースの生成に利用した知識リソースの大きさや、知識リソースに含まれていた情報の偏りに由来するリレーション値の絶対値を、各知識データベース内でのリレーション値の総和に対する相対的な値や、注目したリレーションに対する相対的な値に変換することで、対象となる知識データベースの情報を正しく反映した演算や利用を可能とする。

## 【 0 0 9 3 】

標準化の方法としては、知識データベース生成に用いた知識リソースにおいて抽出した句オブジェクトの出現頻度により標準化する方法がある。たとえば、言語化知識の知識リソースからの知識データベースの生成において、句オブジェクト毎に抽出回数を記録することで、知識データベースの演算や利用の際、対象の各知識データベースにおいて、全句オブジェクトの抽出回数の和、又は注目する句オブジェクトの抽出回数の和を求め、その値が対象の知識データベース間で同じ値に成るように全リレーションの値を標準化する。この方法を知識データベースの和に用いる場合は、演算元の両知識データベースにおいて標準化された句オブジェクトの抽出回数の和を、演算結果の知識データベース内に記録することで、引き続き、他の知識データベースとの和において標準化を行うことが可能となる。

## 【 0 0 9 4 】

また、演算や利用の対象である知識データベースに含まれる句オブジェクト間のリレーション値に注目して標準化する方法がある。たとえば、知識データベースの演算や利用の際、対象の各知識データベースにおいて、各知識データベースに含まれる句オブジェクト間のリレーション値の総和、あるいは、注目する句オブジェクトに関連した句オブジェクトとの間のリレーション値の総和、あるいは、注目する1つのリレーションの値を求め、その値が対象の知識データベース間で同じ値に成るように全リレーションの値を標準化する。

## 【 0 0 9 5 】

注目する句オブジェクトの知識リソースにおける出現頻度や、知識データベー

ス内の注目するリレーションの値によって標準化を行うことにより、知識データベース全体、あるいは注目する句オブジェクトやリレーションの周辺における、知識データベース間の知識の偏りを対象としたより正確な処理が可能となる。

## 【 0 0 9 6 】

知識データベースに含まれるオブジェクトに関する情報の表示において、対象となるオブジェクトを選択する方法としては、知識データベースに含まれるオブジェクトを指定又は検索することで行う。オブジェクトの指定においては、オブジェクトの値を入力する。たとえば、句オブジェクトの指定のための値の入力においては、指定したい各オブジェクトの値の入力に従って、先頭一致で同じ値のオブジェクト値を持つ句オブジェクトが候補としてリスト表示され、そこから選択できるユーザインターフェースであることが望ましい。検索においては、オブジェクトの種類毎に用意した検索方法によって、オブジェクトの値で検索を行うとよい。たとえば、句オブジェクトについては句の正規化表現による検索を、DNA配列オブジェクト、蛋白質配列オブジェクトについては相同性検索による検索を、また、画像オブジェクトについては画像のパターンマッチングによる検索を用意するとよい。指定又は検索により選択された、1つ又は複数のオブジェクトに関して関連する情報をリスト表示又はグラフィックスにて表示する方法を用意するとよい。リスト表示において表示する情報としては、オブジェクトの種類、オブジェクトの値、そのオブジェクトに関連するリレーションのリレーション値の合計、そのオブジェクトと指定した値以上のリレーションを持つオブジェクトの数、そのオブジェクトと同一であるとみなす値を持つオブジェクトの数、また、句オブジェクトの場合は知識リソースからの抽出回数などとする。指定されたオブジェクトが1つの場合には、そのオブジェクトと指定した値以上のリレーションを持つオブジェクトに関して、同様の情報のリスト表示を同時に表示するとよい。グラフィックス表示においては、例えば、対象オブジェクトを示すマーカを円周上に等間隔で並べ、マーカ間でリレーションを示す線を引く。リレーション値の大小を線の太さで表すことにより、対象オブジェクト全体のリレーションの関係を容易に識別可能とする。

## 【 0 0 9 7 】

図 1 9 は、知識データベースにおいて、指定した複数のオブジェクトの関連を表示するグラフィック表示の例を示す図である。知識データベースに含まれる任意のオブジェクトを指定、もしくは、検索によって選択し、円周 1 9 0 1 上にこれらのオブジェクト 1 9 0 2 を等間隔に配置する。そして、全てのオブジェクト間を、リレーションを示す線 1 9 0 3 で結ぶ。リレーション値の大小を線の太さで表し、また、必要に応じて線上にリレーション値を表示することにより、どのオブジェクト間が強いリレーションを持っているかを明示する。また、多数のオブジェクトを同時に表示した場合など、指定した値以上のリレーション値であるリレーションのみを表示することによって表示を簡略化する。

## 【 0 0 9 8 】

本知識データベースの基本的な利用方法としては、調査対象のキーワードを、目的に応じて構築した知識データベースに与え、そのキーワードに関連する情報を表示することがある。その際、目的に応じて構築した知識データベースから情報を抽出する方法としては、指定したキーワードに対応する 1 つ又は複数のオブジェクト（クエリーオブジェクト）に対して、階層的に、強いリレーションを持つオブジェクト（ターゲットオブジェクト）を選択していく方法が有効である。

## 【 0 0 9 9 】

すなわち、第 1 階層として、クエリーオブジェクトに関連したリレーションをリレーション値の高い順にソートし、予め設定した数のリレーションを上位のものから選択し、選択したリレーションで関連付けられたオブジェクト（ターゲットオブジェクト）を得る。次に、第 2 階層として、第 1 階層にて得たターゲットオブジェクトを新たなクエリーオブジェクトとして同様の処理を行い、次のターゲットオブジェクトを得る。指定した回数だけ、階層的にターゲットオブジェクトを得ることにより、第 1 階層において与えたクエリーオブジェクトに階層的に関連したオブジェクトを得ることができる。この方法において、クエリーオブジェクトに関連したリレーションを選択する際、選択するリレーション値の最低値を設定し、その値によって足切りを行うことによって、主要なリレーションのみで構成された、階層的なオブジェクトのネットワークを得ることもできる。

## 【 0 1 0 0 】

また、他の基本的な利用方法としては、複数のキーワードに対して、それらを1つのグループと考え、そのグループに関連する情報を表示する方法もある。その際、目的に応じて構築した知識データベースから情報を抽出する方法としては、指定したキーワードに対応する複数のオブジェクトを最初のクエリーオブジェクトとしたとき、第1階層のターゲットオブジェクト取得に先立ち、全クエリーオブジェクトおよびそれらに関連したリレーションを合併して1つのオブジェクトとし、新たにこれをクエリーオブジェクトとしてから、階層的にオブジェクトを選択していく。この方法によれば、グループに所属するオブジェクトに対して平均して高いリレーション値を持つリレーションにより結び付けられたオブジェクト得ることが可能である。例えば、調査対象のキーワードが多い場合など、キーワードが集合として強く関連を持つオブジェクトを得ることは有用である。

#### 【0101】

上記の方法により得られた階層的にリレーションを持つオブジェクトの表示方法として、テーブル形式のリスト表示と、グラフィカル表示を用意する。グラフィカル表示においては、第1階層として与えたクエリーオブジェクトが1つであった場合、クエリーオブジェクトを中心として、後の階層ほど半径の大きくなる同心円上にターゲットオブジェクトを並べ、クエリーオブジェクトとターゲットオブジェクトとの間を線で結ぶことで、オブジェクト間のリレーションと階層を効果的に表示することができる。このとき、リレーション値の大きさを線の太さで表現するとよい。また、表示上で同じオブジェクトは重複して表示はしない。すなわち、オブジェクトとリレーションの表示は低い階層から高い階層へ行い、より低い階層や同じ階層で既に表示されたオブジェクトについては新たには表示をせず、リレーションも、既に表示済みのオブジェクトとの間で線を結ぶ。

#### 【0102】

また、第1階層として与えたクエリーオブジェクトが1つであろうと、複数であろうと、階層毎にオブジェクトを、水平、あるいは、垂直方向に平行に並べ、クエリーオブジェクトとターゲットオブジェクトとの間を線で結ぶことで、オブジェクト間のリレーションと階層を効果的に表示することができる。このとき、リレーション値の大きさを線の太さで表現する。また、表示上で同じオブジェク

トは重複して表示はしない。すなわち、オブジェクトとリレーションの表示は低い階層から高い階層へ行い、より低い階層や同じ階層で既に表示されているオブジェクトについては新たには表示をせず、リレーションも、既に表示済みのオブジェクトとの間で線を結ぶ。

#### 【0103】

階層的にリレーションを持つオブジェクトのグラフィカル表示において、表示されたオブジェクトから1つ又は複数の任意のオブジェクトを選択状態とすることを可能とするとよい。すなわち、それらを手動でドラッグし再配置することで表示の改善が行えることが望ましい。また、選択したオブジェクトにつき、次の階層で表示するターゲットオブジェクトの数や最低リレーション値の設定を変更し、表示に反映するユーザインターフェースを用意する。また、選択したオブジェクトにつき、それらをクエリーオブジェクトとして、指定した知識データベースより、あらたに知識情報の抽出と表示を行うことを可能にするとよい。また、オブジェクトを複数選択した場合、それらをシノニムオブジェクトグループとして登録し、知識データベースの再構成を行うことを可能にするとよい。その際、選択したオブジェクトの内、どれを代表オブジェクトとするか指定することを可能とする。

#### 【0104】

図20は、知識データベースにおいて、指定した1つのオブジェクト（クエリーオブジェクト）に、階層的に関連したオブジェクトと、それら相互の間のリレーションを同心円上に広がる形で表示するグラフィックス表示インターフェースの例を示す図である。

#### 【0105】

クエリーオブジェクト2001を中心に表示し、そのクエリーオブジェクト2001に対して、指定した以上に高いリレーションを持つオブジェクト2002を指定した数だけ知識データベースより抽出してこれを第1階層のオブジェクトとし、クエリーオブジェクトを中心とした円周2004上に等間隔に配置する。続いて、第1階層の各オブジェクトに対して、それぞれ、指定した以上の高いリレーションを持つオブジェクト2003を指定した数だけ知識データベースより

抽出し、これを第2階層のオブジェクトとして、クエリーオブジェクトを中心とする第1階層の円2004よりも半径の大きな円2005上に等間隔に配置する。この際、第1階層までにすでに表示されているオブジェクトについては、第2階層には含めない。また、各オブジェクトは、できるだけそれらが高いリレーションを持つ第1層のオブジェクトの近くに配置するようにする。このようにして、指定した階層だけ、階層的にオブジェクトを抽出して表示していく。

#### 【0106】

また、表示した全てのオブジェクト間をリレーションを表す線2006で結ぶ。リレーション値の大小を線の太さで表し、また、必要に応じて線上にリレーション値を表示することにより、どのオブジェクト間が強いリレーションを持っているかを明示する。多数のオブジェクトを同時に表示した場合など、指定した値以上のリレーションを示す線のみを表示することによって表示を簡略化してもよい。表示されたオブジェクトは選択可能であり、選択したオブジェクトはドラッグして、見やすい位置に移動することができるのが望ましい。オブジェクトを移動した際はリレーションを示す線も同時に移動する。選択したオブジェクトについて、次の階層に表示するオブジェクトの数を増減したり、選択したオブジェクト自身を表示より削除したりすることができるのが望ましい。また、選択したオブジェクトにつき、それらをクエリーオブジェクトとして、指定した知識データベースより、あらたに知識情報の抽出と表示を行うことを可能にするとよい。

#### 【0107】

図21は、知識データベースにおいて、指定した1つ又は複数のオブジェクトに、階層的に関連したオブジェクトと、それら相互の間のリレーションを平行に広がる形で表示するグラフィックス表示の例を示す図である。

クエリーオブジェクト2101を直線状に、等間隔に表示する。表示においては、多数のオブジェクトを表示した際のみやすさを改善する方法として、互い違いになるように表示してもよい。各クエリーオブジェクトに対して、それぞれ、指定した以上に高いリレーションを持つオブジェクト2102を指定した数だけ知識データベースより抽出してこれを第1階層のオブジェクトとし、クエリーオブジェクトに平行な位置2104に等間隔に配置する。続いて、第1階層の各オ

ブジェクトに対して、それぞれ、指定した以上の高いリレーションを持つオブジェクト2103を指定した数だけ知識データベースより抽出し、これを第2階層のオブジェクトとして、クエリーオブジェクトに平行な第1階層よりも離れた位置2105に配置する。この際、第1階層までにすでに表示されているオブジェクトについては、第2階層には含めない。また、各オブジェクトは、できるだけそれらが高いリレーションを持つ第1層のオブジェクトの近くに配置するようにする。

#### 【0108】

このようにして、指定した階層だけ、階層的にオブジェクトを抽出して表示していく。また、全てのオブジェクト間を、リレーションを表す線2106で結ぶ。リレーション値の大小を線の太さで表し、また、必要に応じて線上にリレーションの値を表示することにより、どのオブジェクト間が強いリレーションを持っているかを明示する。多数のオブジェクトを同時に表示した場合など、指定した値以上のリレーションのみを表示することによって表示を簡略化してもよい。表示されたオブジェクトは選択可能であり、選択したオブジェクトはドラッグして、見やすい位置に移動することができるのが望ましい。オブジェクトを移動した際はリレーションを示す線も同時に移動する。選択したオブジェクトについて、次の階層に表示するオブジェクトの数を増減したり、選択したオブジェクト自身を表示より削除したりすることができるのが望ましい。また、選択したオブジェクトにつき、それらをクエリーオブジェクトとして、指定した知識データベースより、あらたに知識情報の抽出と表示を行うことを可能にするとよい。

#### 【0109】

本知識データベースの応用的な利用方法としては、医学・生物学関連データベースの検索の結果得られる情報を調査対象とし、対応するオブジェクトをクエリーオブジェクトとして、知識データベースから情報を引き出す方法がある。その際、医学・生物学関連データベースの検索の結果得られる情報としては、専門用語、分子名、遺伝子名、生物種、URLなどを記述したの句、DNA配列、蛋白質配列などの分子構造、生体組織画像などの画像、などがある。たとえば、DNA配列データベースや蛋白質配列データベースに対するDNA配列や蛋白質配列の相同性検

索の結果得られる情報のうち、アクセッション番号、DNA配列名、蛋白質名などは、句オブジェクトを含む知識データベースに与え、同じオブジェクト値を持つ句オブジェクトをクエリーオブジェクトとして、得られる情報を表示する。又は、相同性検索の結果得られる情報のうち、DNA配列や蛋白質配列は、DNA配列オブジェクトや蛋白質配列オブジェクトを含む知識データベースに与え、相同性の高いオブジェクト値を持つDNA配列オブジェクトや蛋白質配列オブジェクトをクエリーオブジェクトとして、得られる情報を表示する。

【0110】

また、医学・生物学実験や、データ解析の結果得られる情報を調査対象とし、対応するオブジェクトをクエリーオブジェクトとして、知識データベースから情報を引き出す方法がある。その際、医学・生物学関連データベースの検索の結果得られる情報としては、専門用語、分子名、遺伝子名、生物種、URLなどを記述したの句、DNA配列、蛋白質配列などの分子構造、生体組織画像などの画像、などがある。たとえばDNA配列決定実験においては、DNAシーケンサにより決定されたDNA配列を、DNA配列オブジェクトや蛋白質配列オブジェクトを含む知識データベースに与え、相同性の高いオブジェクト値を持つDNA配列オブジェクトや蛋白質配列オブジェクトをクエリーオブジェクトとして、得られる情報を表示する。また、遺伝子発現データのクラスタ解析の結果得られる遺伝子名のクラスタを、句オブジェクトを含む知識データベースに与え、各遺伝子名と同じオブジェクト値を持つ句オブジェクトをクエリーオブジェクトとして、得られる情報を表示する。

【0111】

図22は、遺伝子の発現プロファイル解析の結果得られた、遺伝子名と同じオブジェクト値を持つ句オブジェクトのクラスタをクエリーオブジェクトとして、図21に示した表示方法により、階層的に関連したオブジェクトの情報を表示する、知識データベースの応用例を示す図である。

【0112】

遺伝子の発現プロファイルの解析において、クラスタ解析アプリケーション201によって分類された遺伝子のクラスタを選択(2203)し、それらを知

識データベースアプリケーション 2 2 0 2 に渡し、同じオブジェクト値を持つ句オブジェクトをクエリーオブジェクト 2 2 0 4 として、指定した知識データベースより階層的にオブジェクトとリレーションを抽出し、図 2 1 と同様の表示方法で表示したものである。このように、知識データベースの応用によって、医学、生物学実験やデータ解析の結果より、既存の知識リソースに存在する知識を得ることができる。

【 0 1 1 3 】

【発明の効果】

以上説明したように、本発明によれば、医学・生物学の知識を構造化して蓄積・利用することが可能となり、知識の説明、知識の表現、知識の発見、情報システムによる知識の利用が可能となる。

【図面の簡単な説明】

【図 1】

本発明のシステム構成図。

【図 2】

最も単純な知識データベースの概念図。

【図 3】

知識データベースの概念図。

【図 4】

知識データベースの合併の概念図。

【図 5】

知識データベース内におけるオブジェクトとリレーションの合併の概念図。

【図 6】

異なる種類のオブジェクトを含む知識データベースの例を示す図。

【図 7】

リレーション値によるオブジェクト・リレーション抽出の説明図。

【図 8】

一般的な句オブジェクト切り出しとリレーション設定の手順の一例を示すフローチャート。

【図 9】

句オブジェクト切り出しとリレーション設定の例を示す説明図。

【図 1 0】

索引からの句オブジェクト切り出しとリレーション設定の手順の一例を示すフローチャート。

【図 1 1】

索引からの句オブジェクト切り出しとリレーション設定の例を示す説明図。

【図 1 2】

DNAデータベースエントリからのDNA配列オブジェクトを含む知識データベース生成の例を示す図。

【図 1 3】

画像ファイルと画像の説明からの画像オブジェクトを含む知識データベース生成の例を示す図。

【図 1 4】

Webページからの知識データベース生成の例を示す図。

【図 1 5】

知識データベースデータ構造の例を示す図。

【図 1 6】

知識データベースにおけるオブジェクト値、リレーション値の例を示す図。

【図 1 7】

知識データベースデータ構造でのデータ格納の例を示す図。

【図 1 8】

知識データベース比較グラフの例を示す図。

【図 1 9】

知識データベースに含まれるオブジェクト・リレーション表示の例を示す図。

【図 2 0】

関連オブジェクトの階層的表示（円形表示）の例を示す図。

【図 2 1】

関連オブジェクトの階層的表示（平行表示）の例を示す図。

【図22】

クラスタ解析からの知識データベース情報検索の例を示す図。

【符号の説明】

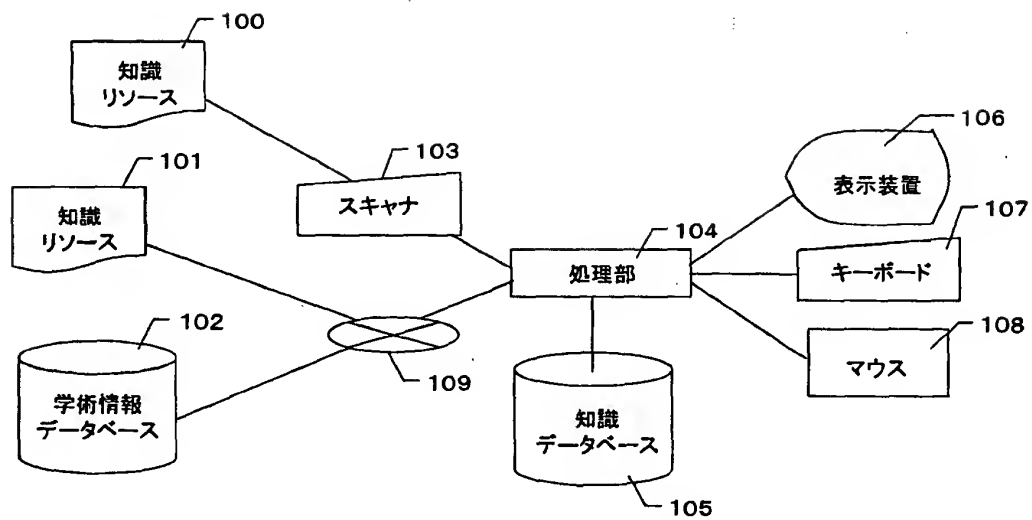
100…教科書、専門書など印刷された医学・生物学の言語化知識データ、101…言語化知識データ、102…学術情報データベース、103…スキャナ装置、104…処理部、105…データベース、106…表示装置、107…キーボード、108…マウス、201…オブジェクト、202…リレーション、203…最も単純な知識データベース、301…知識データベース、401…知識データベースA、402…知識データベースB、403…知識データベースAB、404…オブジェクトa、405…オブジェクトb、406…オブジェクトc、407…オブジェクトd、408…オブジェクトe、409…合併した知識データベースの両方にもともと存在していたリレーション、410…合併した知識データベースの一方にのみ存在していたリレーション、411…合併により新たに発生したリレーション、501…オブジェクト合併前の知識データベース、502…オブジェクト合併後の知識データベース、601…異なる種類のオブジェクトが存在する知識データベース、602-606…句オブジェクト、607-609…DNA配列オブジェクト、610…句オブジェクト間のリレーション、611…DNA配列オブジェクト間のリレーション、612-614…句オブジェクトとDNA配列オブジェクトの間で、同一であるとみなす値を持つリレーション、701…高いリレーションを持つオブジェクトを含む知識データベース、702…異なる種類のオブジェクト間のリレーションから求められたリレーション、801-816…フローチャートの処理、901…論文のタイトル、902…切り出した論文のタイトル、903…句、904…句分離文字列、905…句オブジェクト、906…句オブジェクト間のリレーション、1001-1014…フローチャートの処理、1101…索引、1102…索引に含まれる句、1103…索引に含まれる頁番号、1104…頁出現パターンデータ、1105…知識データベース（オブジェクトとリレーション）、1106…知識データベースの図、1201…DNAデータベースエントリ、1202…DNA配列に付随する言語化知識、1203…DNA配列の識別子、1204…DNA配列、1205…句、1206…句分離文字列

、 1 2 0 7 … 句オブジェクト、 1 2 0 8 … DNA配列に付随する言語化知識から生成した句オブジェクトとDNA配列の識別子から生成した句オブジェクトとの間のリレーション、 1 2 0 9 … DNA配列の識別子から生成した句オブジェクト、 1 2 1 0 … DNA配列オブジェクト、 1 2 1 1 … DNA配列オブジェクトとDNA配列の識別子から生成した句オブジェクトとの間のリレーション、 1 3 2 1 … 画像ファイル、 1 3 2 2 … 画像、 1 3 2 3 … 画像の識別子（画像ファイル名）、 1 3 2 4 … 画像の説明、 1 3 2 5 … 画像の識別子（画像ファイル名）、 1 3 2 6 … 画像を説明する言語化知識、 1 3 2 7 … 句、 1 3 2 8 … 句分離文字列、 1 3 2 9 … 句オブジェクト、 1 3 3 0 … 画像を説明する言語化知識から生成した句オブジェクトと画像の識別子から生成した句オブジェクトとの間のリレーション、 1 3 3 1 … 画像の識別子から生成した句オブジェクト、 1 3 3 2 … 画像オブジェクト、 1 3 3 3 … 画像オブジェクトと画像の識別子から生成した句オブジェクトとの間のリレーション、 1 4 4 1 … Webページ、 1 4 4 2 … WebページのURL、 1 4 4 3 … Webページに含まれる言語化知識、 1 4 4 4 … Webページに含まれる画像、 1 4 4 5 … Webページに含まれる画像のURL、 1 4 4 6 … WebページのURLから生成した句オブジェクト、 1 4 4 7 … 句、 1 4 4 8 … 句分離文字列、 1 4 4 9 … 句オブジェクト、 1 4 5 0 … Webページの言語化知識から生成した句オブジェクト及び画像のURLから生成した句オブジェクトとWebページのURLから生成した句オブジェクトとの間のリレーション、 1 4 5 1 … 画像オブジェクト、 1 4 5 2 … 画像のURLから生成した句オブジェクト、 1 4 5 3 … 画像オブジェクトと画像のURLから生成した句オブジェクトとの間のリレーション、 1 5 0 1 … 知識データベーステーブル、 1 5 0 2 … オブジェクトテーブル、 1 5 0 3 … リレーションテーブル、 1 5 0 4 … シノニムオブジェクト情報テーブル、 1 5 0 5 … シノニムオブジェクト辞書テーブル、 1 5 0 6 … 知識データベースを格納するデータベース構造、 1 6 0 1 … 知識データベース「Sample NDB」、 1 6 0 2 … オブジェクトデータ、 1 6 0 3 … リレーションデータ、 1 7 0 1 … 知識データテーブルのレコード、 1 7 0 2 … オブジェクトテーブルのレコード、 1 7 0 3 … リレーションテーブルのレコード、 1 8 0 1 … 知識データベース比較グラフ、 1 8 0 2 … X軸（リレーション値）、 1 8 0 3 … Y軸（知識データベース）、 1 8 0 4 … 折れ線グラフ、 1 8 0 5 … クエ

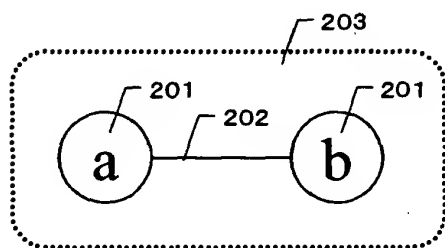
リーオブジェクトのオブジェクト値、1806…グラフの凡例（ターゲットオブジェクトのオブジェクト値）、1901…オブジェクトを表示する円、1902…オブジェクト、1903…リレーション、2001…クエリーオブジェクト、2002…第1階層のオブジェクト、2003…第2階層のオブジェクト、2004…第1階層のオブジェクトを表示する円、2005…第2階層のオブジェクトを表示する円、2006…リレーション、2101…クエリーオブジェクト、2102…第1階層のオブジェクト、2103…第2階層のオブジェクト、2104…第1階層のオブジェクトの表示域、2105…第2階層のオブジェクトの表示域、2106…リレーション、2201…遺伝子発現プロファイルのクラスタ解析結果表示、2202…知識データベースの階層的オブジェクト関連表示、2203…遺伝子発現プロファイルのクラスタ解析結果表示における処理対象遺伝子名の選択、2204…知識データベースの階層的オブジェクト関連表示に渡された遺伝子名に対応したクエリーオブジェクト

【書類名】 図面

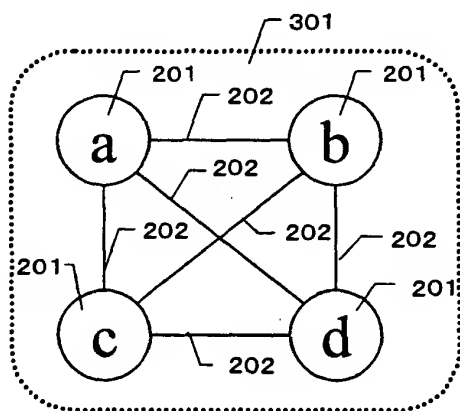
【図1】



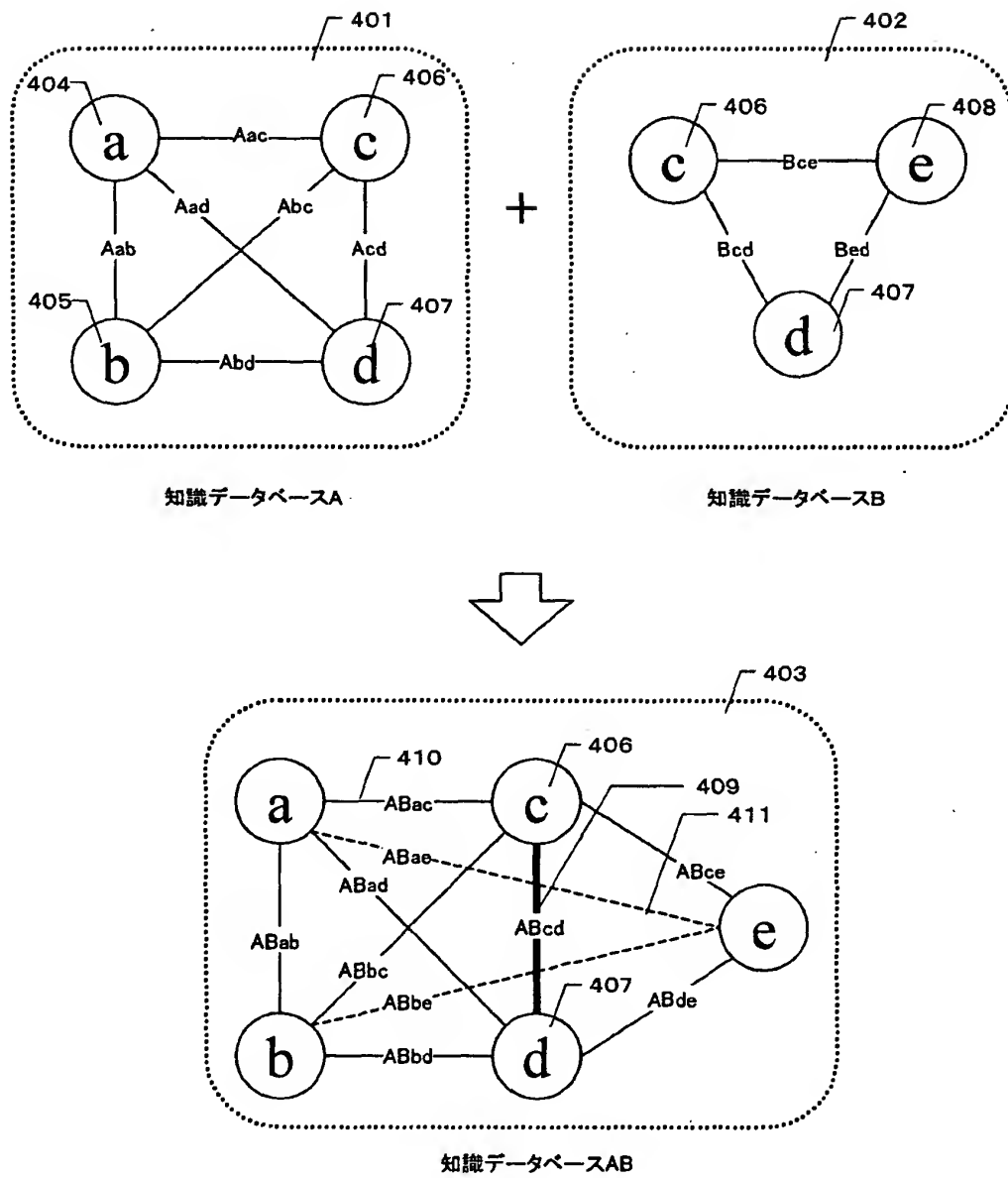
【図2】



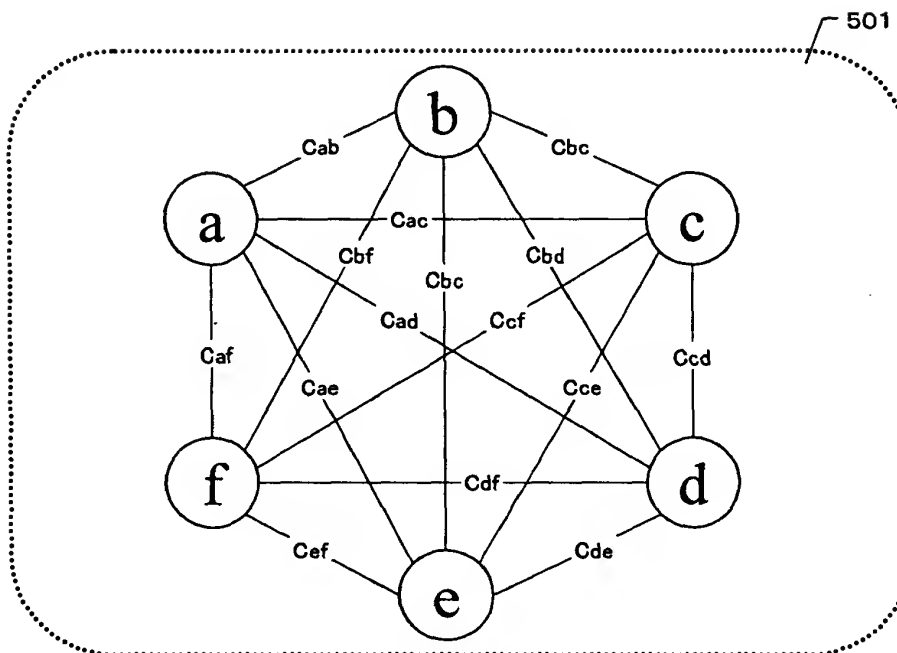
【図3】



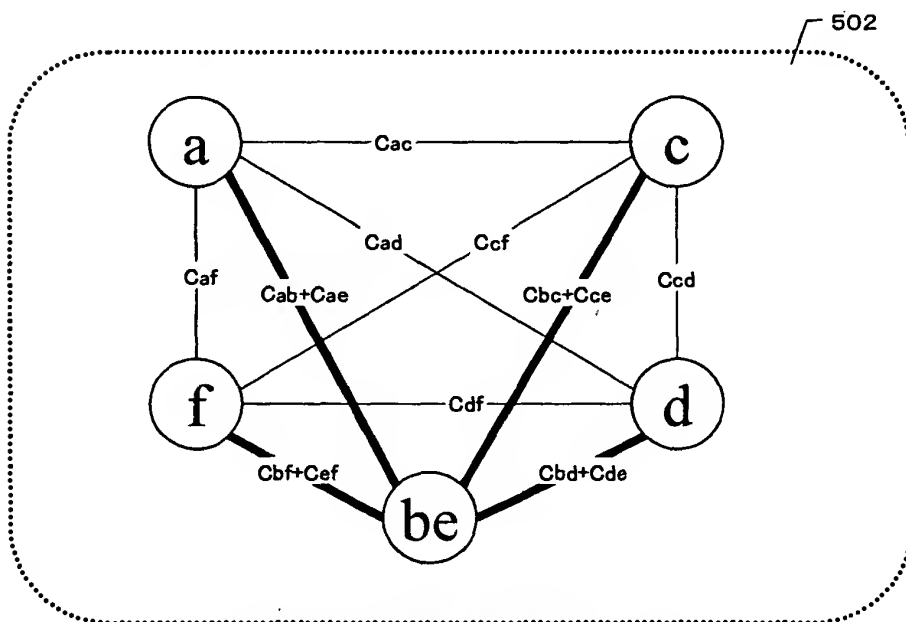
【図4】



【図 5】

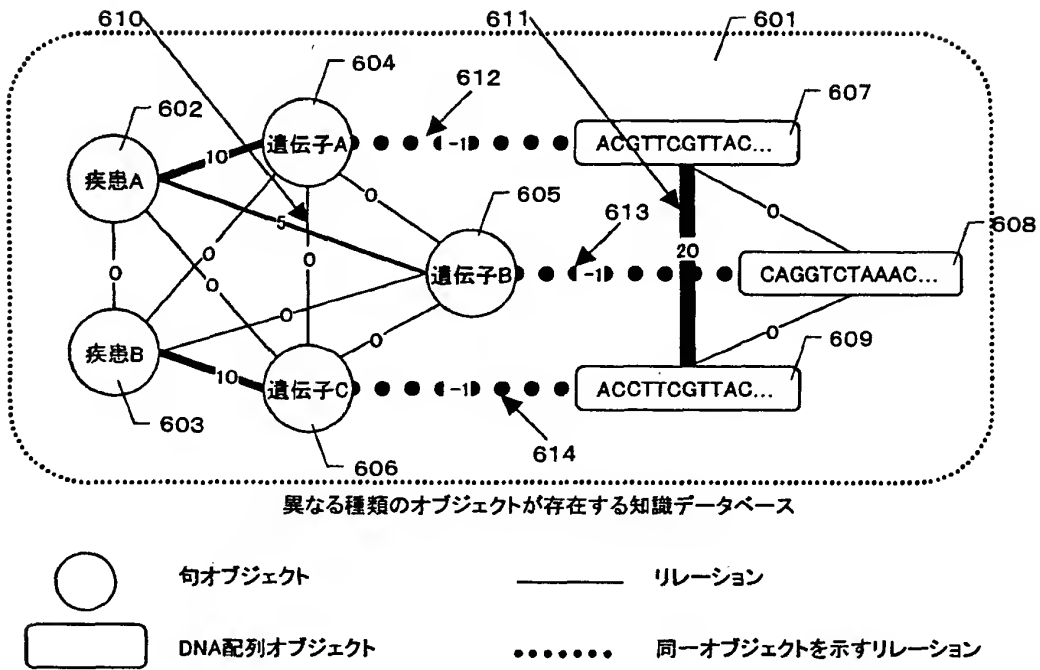


オリジナルの知識データベース

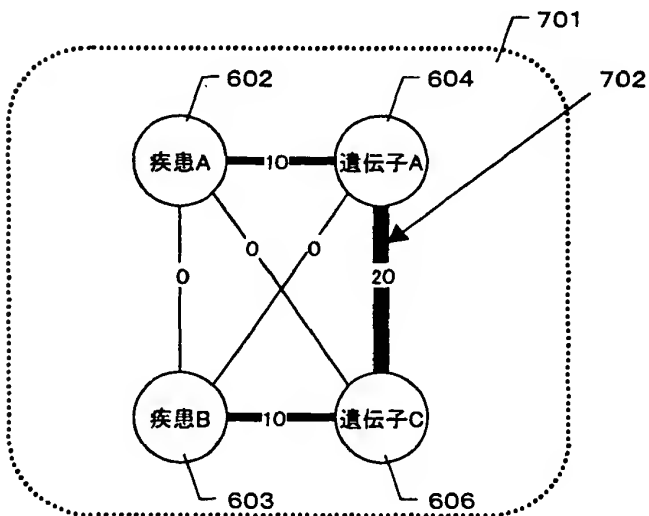


### オブジェクトb、e合併後の知識データベース

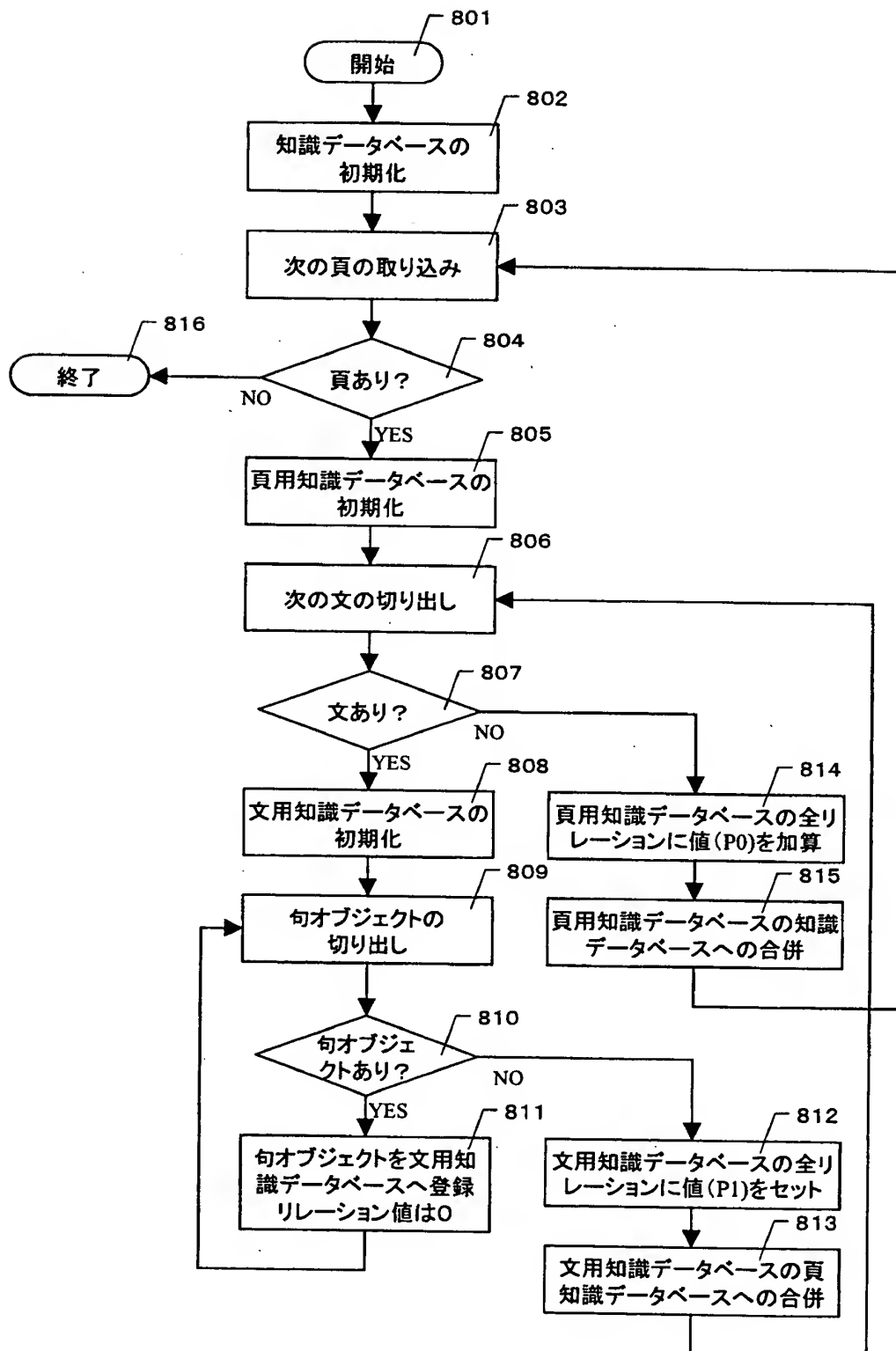
【図 6】



【図 7】

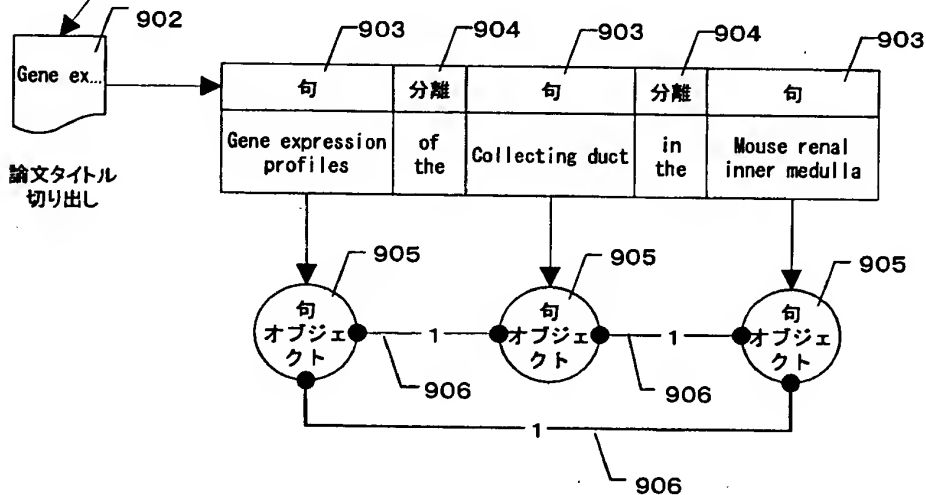


【図8】

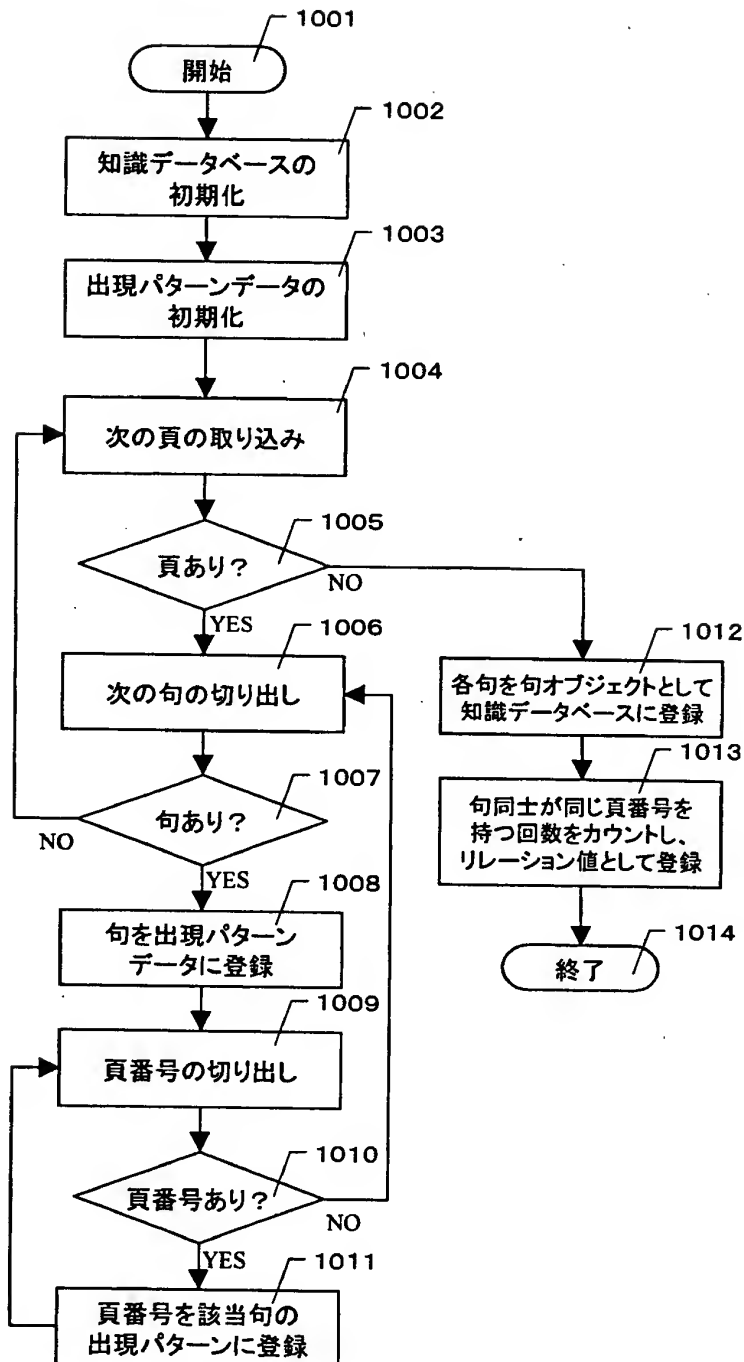


【図 9】

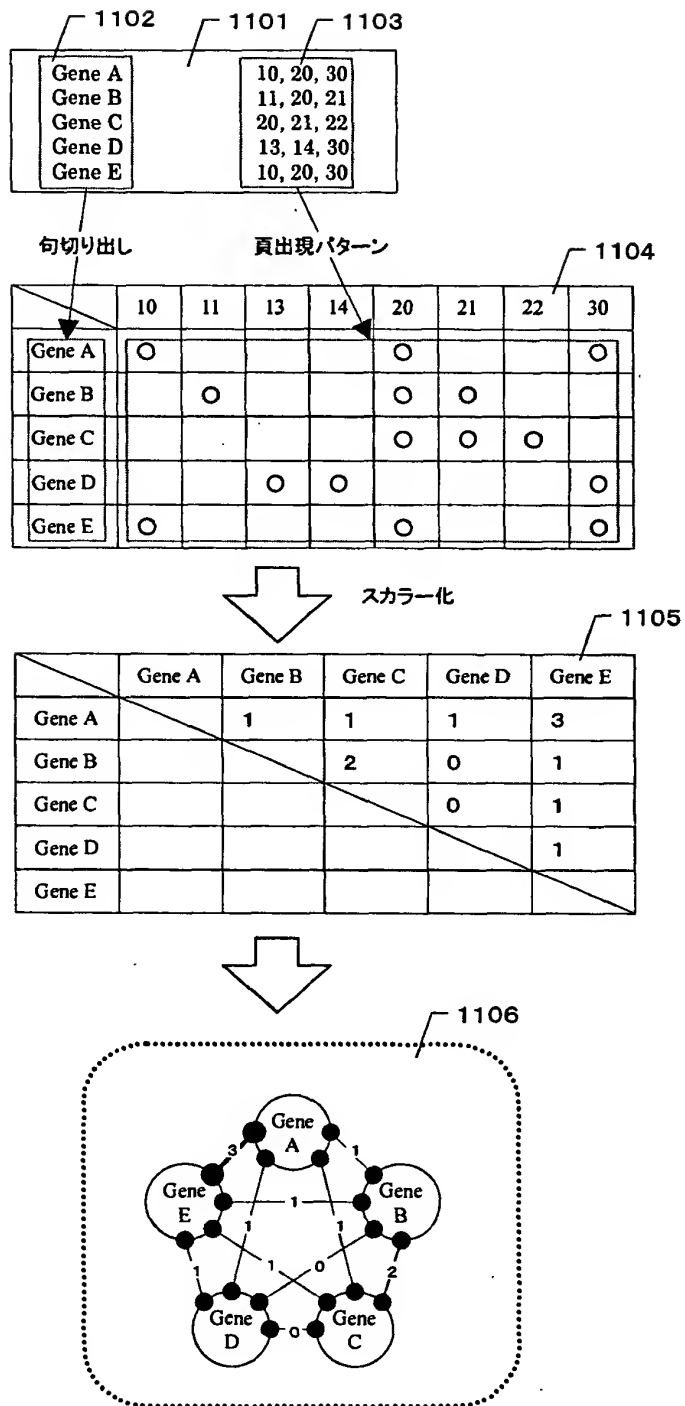
<p><u>Sese J, Nikaidou H, Kawamoto S, Minesaki Y, Morishita S, Okubo K.</u>  BodyMap incorporated PCR-based expression profiling data and a gene ranking system.  Nucleic Acids Res. 2001 Jan 1;29(1):156-158.  [Record as supplied by publisher]  PMID: 11125076</p>	Related Articles
<p><u>Takenaka M, Imai E, Nagasawa Y, Matsuoka Y, Moriyama T, Kaneko T, Hori M, Kawamoto S, Okubo K.</u>  Gene expression profiles of the collecting duct in the mouse renal inner medulla.  Kidney Int. 2000 Jan;57(1):19-24.  PMID: 10620183; UI: 20087124</p>	Related Articles
<p><u>Kawamoto S, Ohnishi T, Kita H, Chisaka O, Okubo K.</u>  Expression profiling by iAFLP: A PCR-based method for genome-wide gene expression profiling  Genome Res. 1999 Dec;9(12):1305-12.  PMID: 10613853; UI: 20082977</p>	Related Articles
<p><u>Hishiki T, Kawamoto S, Morishita S, Okubo K.</u>  BodyMap: a human and mouse gene expression database.  Nucleic Acids Res. 2000 Jan 1;28(1):136-8.  PMID: 10592203; UI: 20063237</p>	Related Articles



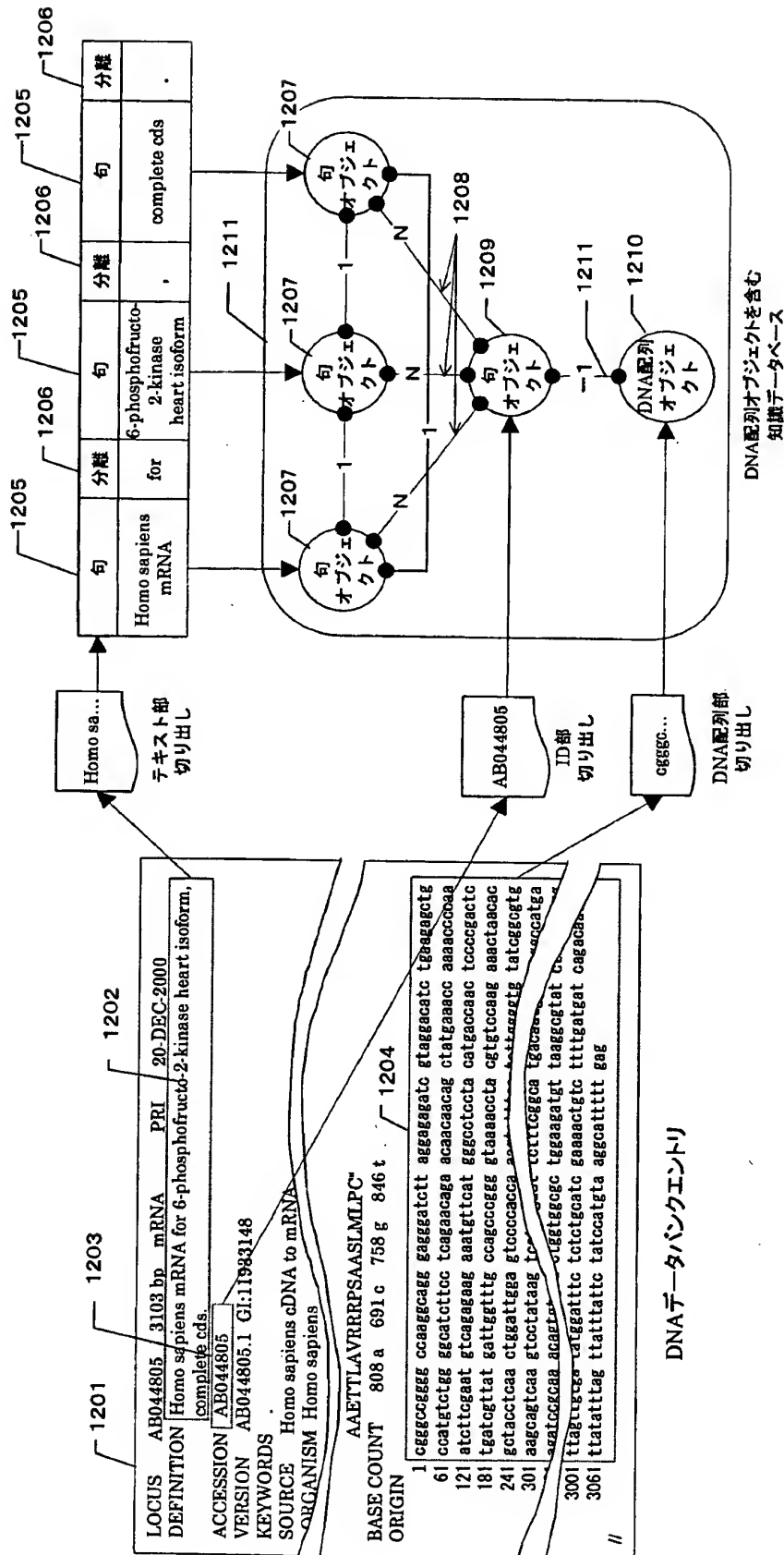
【図10】



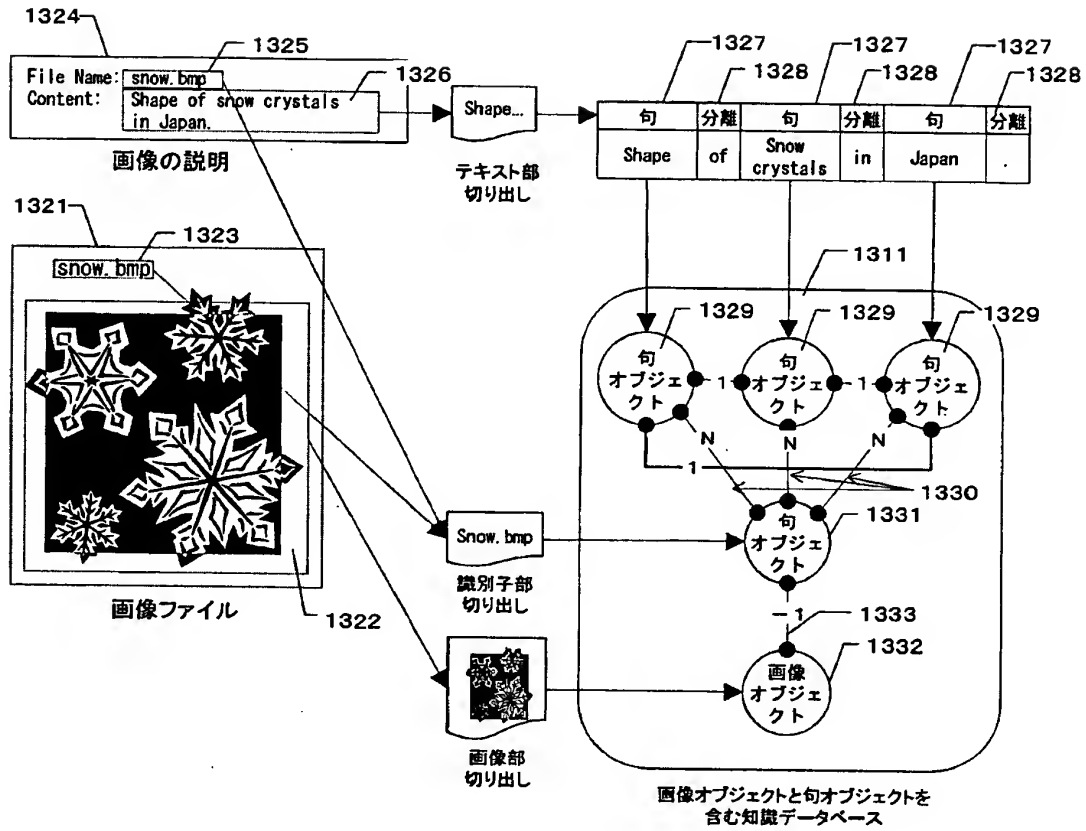
【図 1 1】



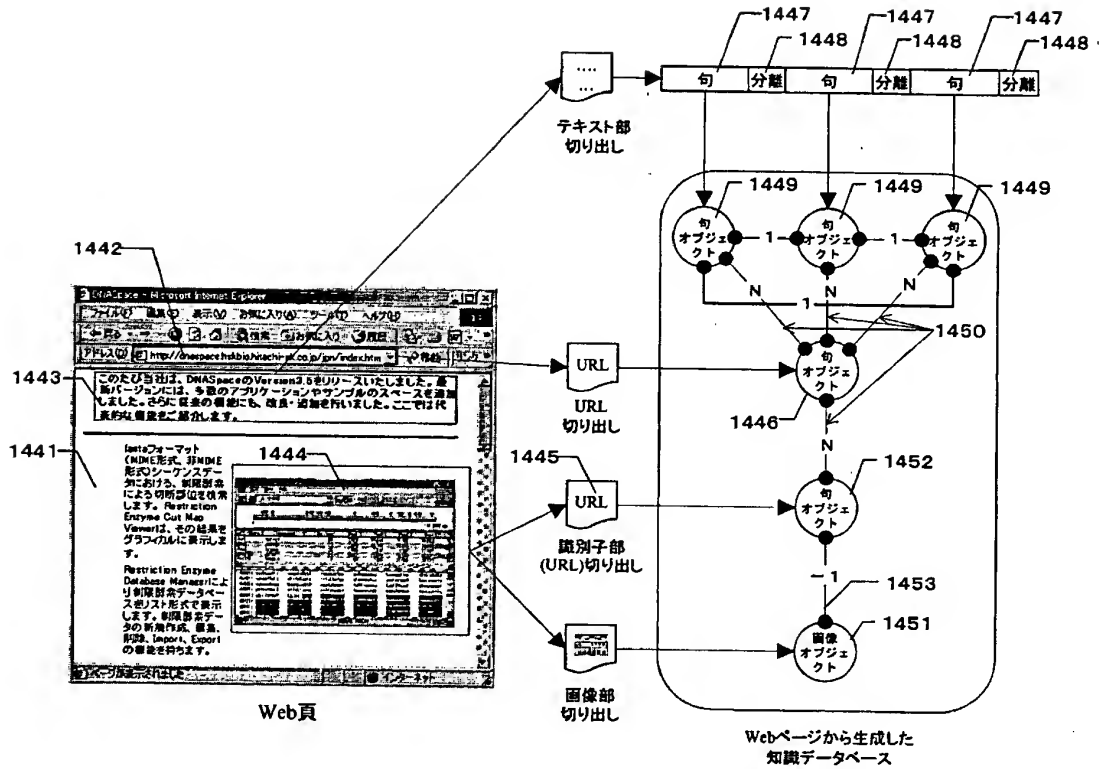
【図 1 2】



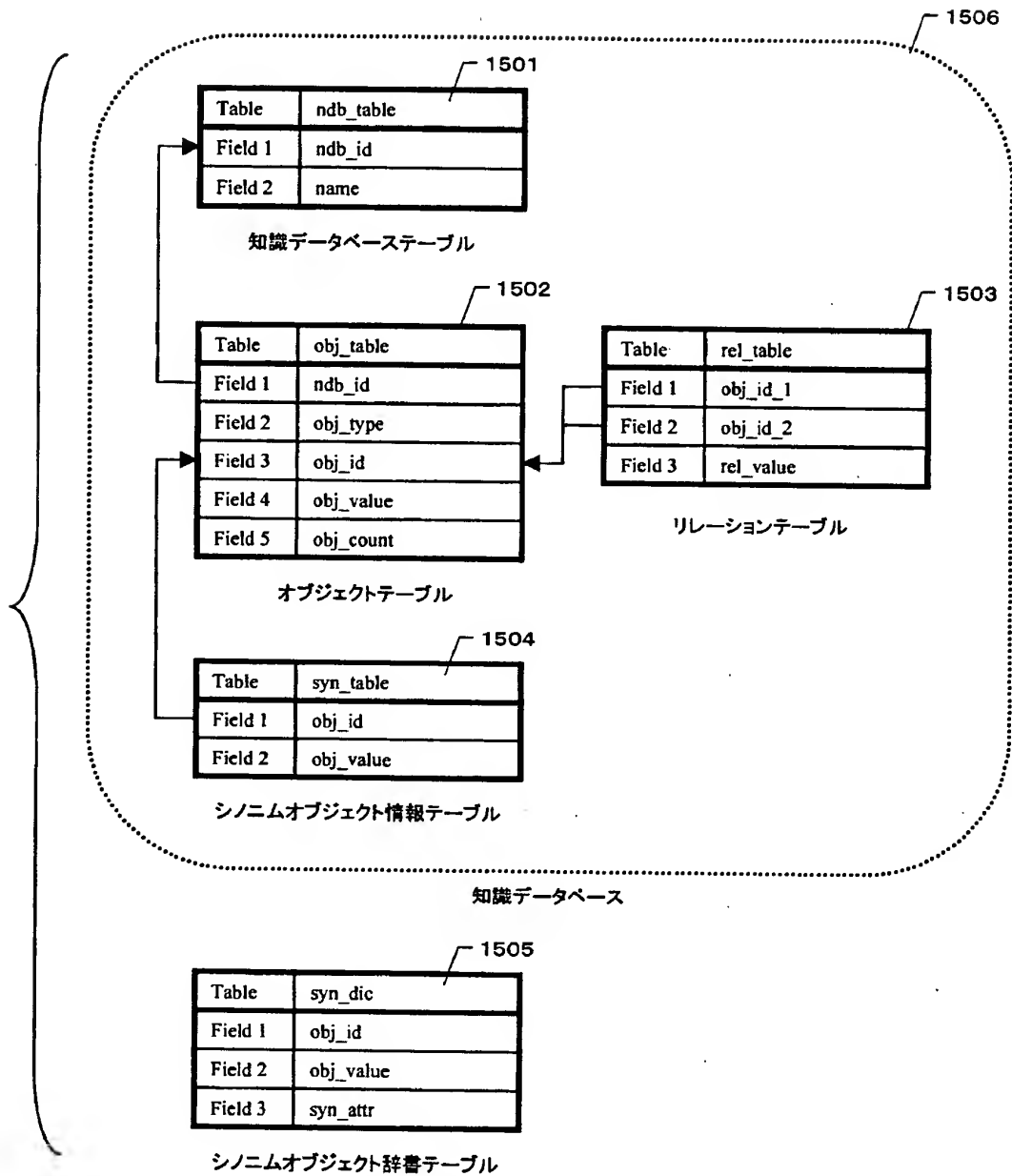
【図 13】



【図14】



【図 15】



【図16】



【図 17】

1701

ndb_id	name
{0701B3FE-775E-4BD0-A851-254397F47B98}	Sample NDB

知識データベーステーブル

1702

ndb_id	obj_id	obj_type	obj_value	obj_count
{0701B3FE-775E-4BD0-A851-254397F47B98}	{0701B3FE-775E-4BD1-A851-254397F47B98}	Phrase	疾患A	15
{0701B3FE-775E-4BD0-A851-254397F47B98}	{0701B3FE-775E-4BD2-A851-254397F47B98}	Phrase	疾患B	10
{0701B3FE-775E-4BD0-A851-254397F47B98}	{0701B3FE-775E-4BD3-A851-254397F47B98}	Phrase	遺伝子A	10
{0701B3FE-775E-4BD0-A851-254397F47B98}	{0701B3FE-775E-4BD4-A851-254397F47B98}	Phrase	遺伝子B	5
{0701B3FE-775E-4BD0-A851-254397F47B98}	{0701B3FE-775E-4BD5-A851-254397F47B98}	Phrase	遺伝子C	10
{0701B3FE-775E-4BD0-A851-254397F47B98}	{0701B3FE-775E-4BD6-A851-254397F47B98}	DNA	ACGTTGTTAC...	null
{0701B3FE-775E-4BD0-A851-254397F47B98}	{0701B3FE-775E-4BD7-A851-254397F47B98}	DNA	CAGGTCTAAAC...	null
{0701B3FE-775E-4BD0-A851-254397F47B98}	{0701B3FE-775E-4BD8-A851-254397F47B98}	DNA	ACCTTCGTTAC...	null

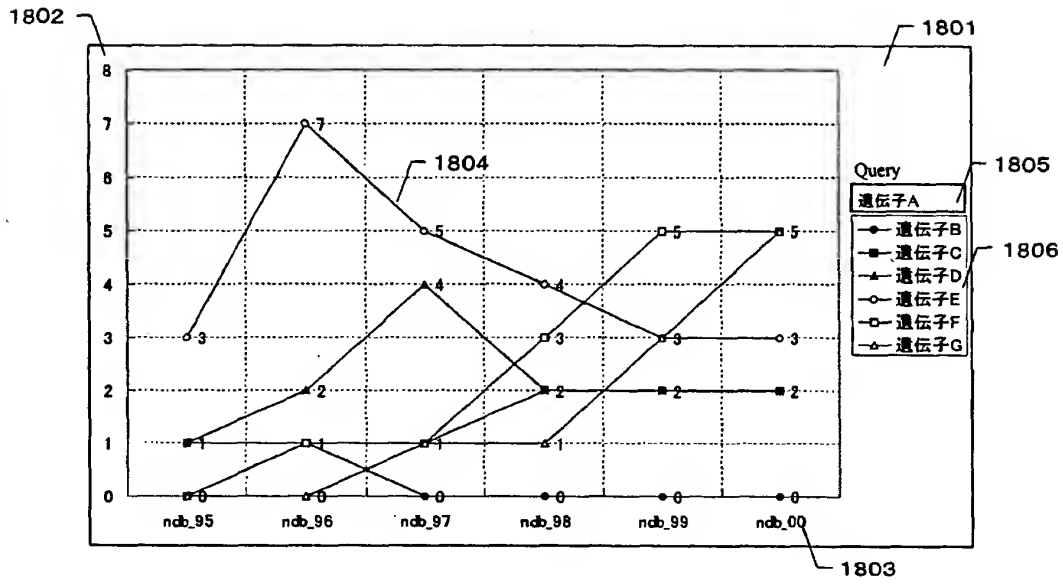
オブジェクトテーブル

1703

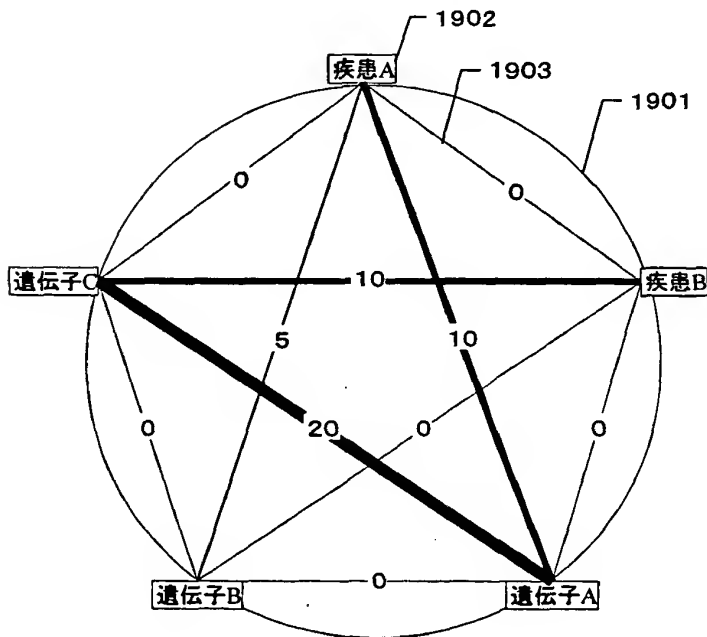
obj_id_1	obj_id_2	rel_value
{0701B3FE-775E-4BD1-A851-254397F47B98}	{0701B3FE-775E-4BD3-A851-254397F47B98}	10
{0701B3FE-775E-4BD1-A851-254397F47B98}	{0701B3FE-775E-4BD4-A851-254397F47B98}	10
{0701B3FE-775E-4BD2-A851-254397F47B98}	{0701B3FE-775E-4BD5-A851-254397F47B98}	10
{0701B3FE-775E-4BD3-A851-254397F47B98}	{0701B3FE-775E-4BD6-A851-254397F47B98}	-1
{0701B3FE-775E-4BD4-A851-254397F47B98}	{0701B3FE-775E-4BD7-A851-254397F47B98}	-1
{0701B3FE-775E-4BD5-A851-254397F47B98}	{0701B3FE-775E-4BD8-A851-254397F47B98}	-1
{0701B3FE-775E-4BD6-A851-254397F47B98}	{0701B3FE-775E-4BD7-A851-254397F47B98}	1
{0701B3FE-775E-4BD6-A851-254397F47B98}	{0701B3FE-775E-4BD8-A851-254397F47B98}	10
{0701B3FE-775E-4BD7-A851-254397F47B98}	{0701B3FE-775E-4BD8-A851-254397F47B98}	1

リレーションテーブル

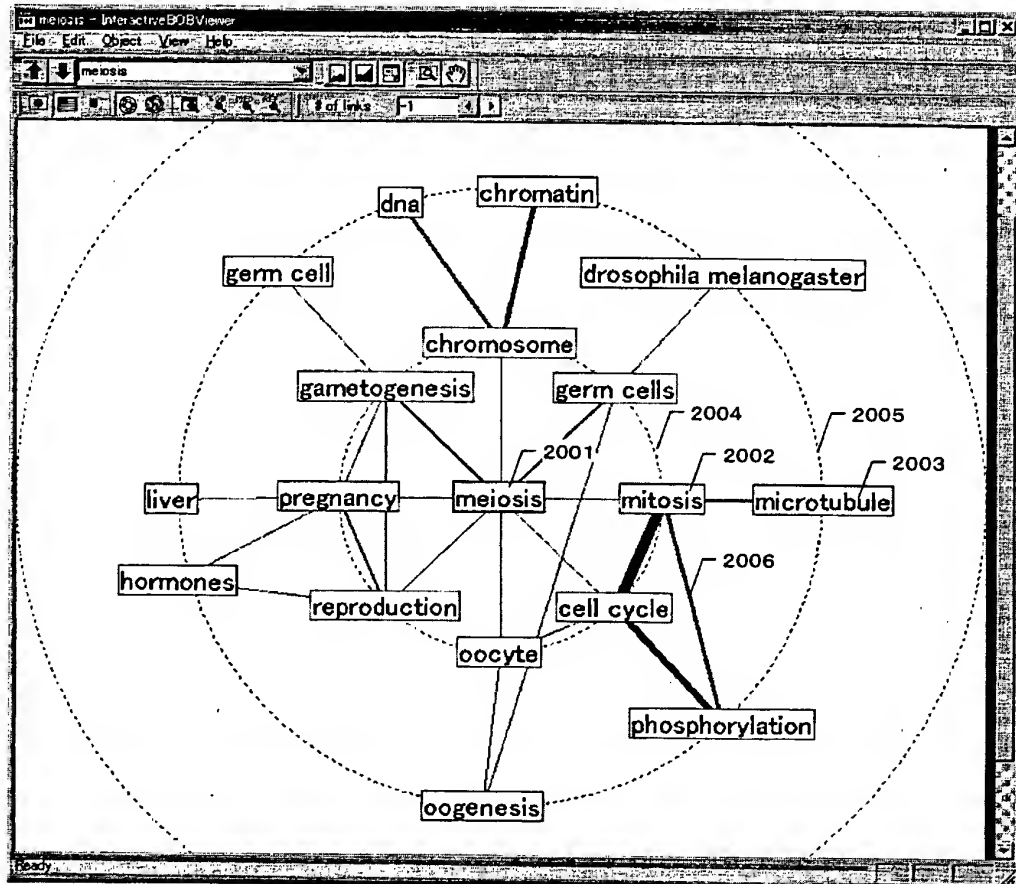
【図18】



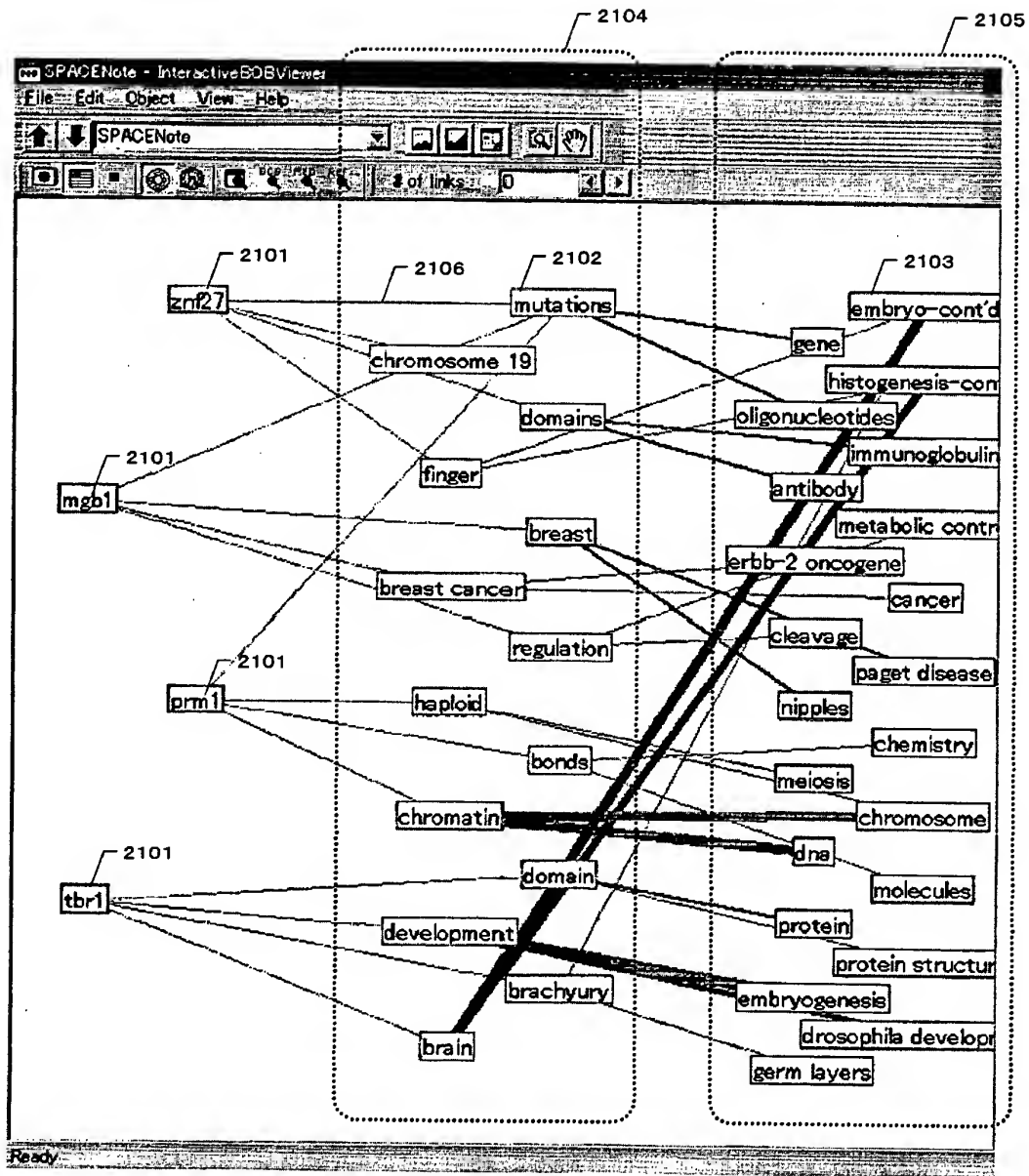
【図19】



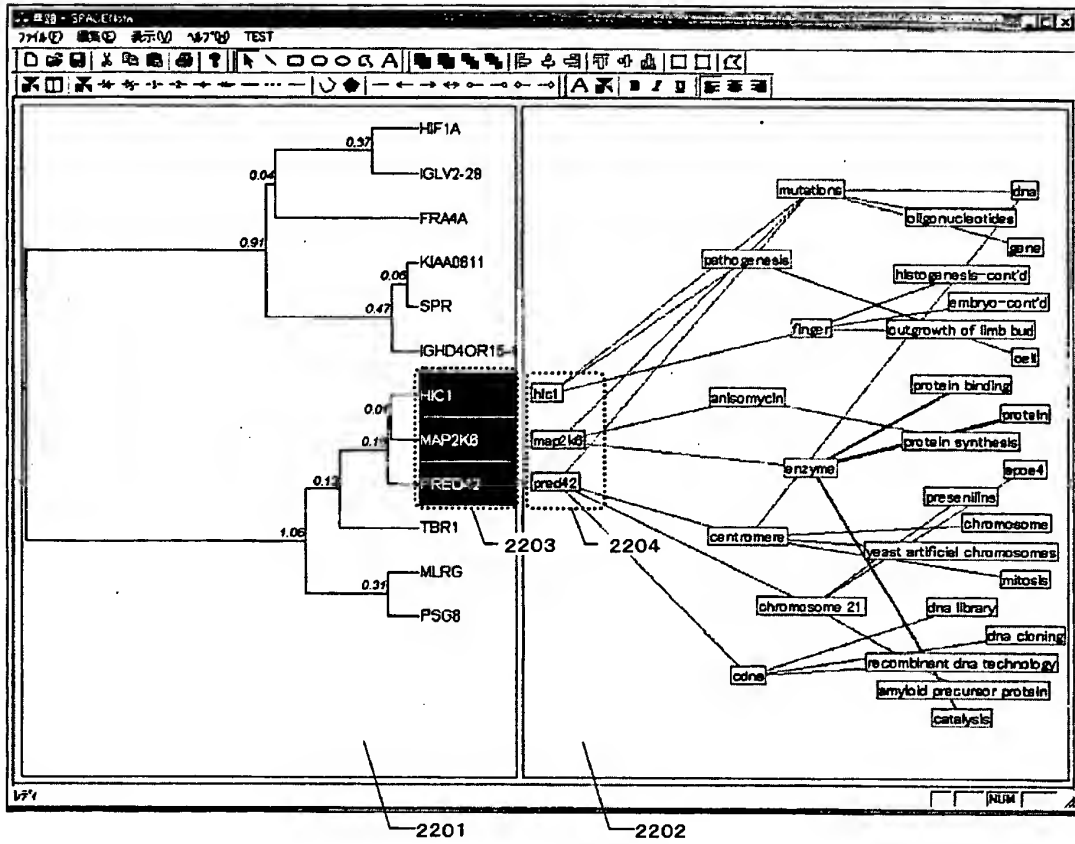
【図20】



【図21】



【图22】



【書類名】 要約書

【要約】

【課題】 医学・生物学の分野において、専門分野を越えた横断的な知識の獲得を可能にする。

【解決手段】 医学分野及び／又は生物学分野で用いられる用語を含む句をオブジェクト値として有するオブジェクトと、オブジェクト相互間の関係を定量的に表現した値をリレーション値として有するオブジェクト間のリレーションとを蓄積した知識データベースから、クエリーオブジェクト2001に関連するオブジェクトをオブジェクト間のリレーションを元に抽出し、クエリーオブジェクト及び抽出されたオブジェクトをその間のリレーションと共に表示する。

【選択図】 図20

特2001-071769

出 願 人 履 歴 情 報

識別番号 [000233055]

1. 変更年月日 1990年 8月 7日

[変更理由] 新規登録

住 所 神奈川県横浜市中区尾上町6丁目81番地

氏 名 日立ソフトウェアエンジニアリング株式会社

特 2001-071769

出 願 人 履 歴 情 報

識別番号 [594082280]

1. 変更年月日 1993年11月12日

[変更理由] 新規登録

住 所 大阪府箕面市瀬川2-11-26

氏 名 大久保 公策